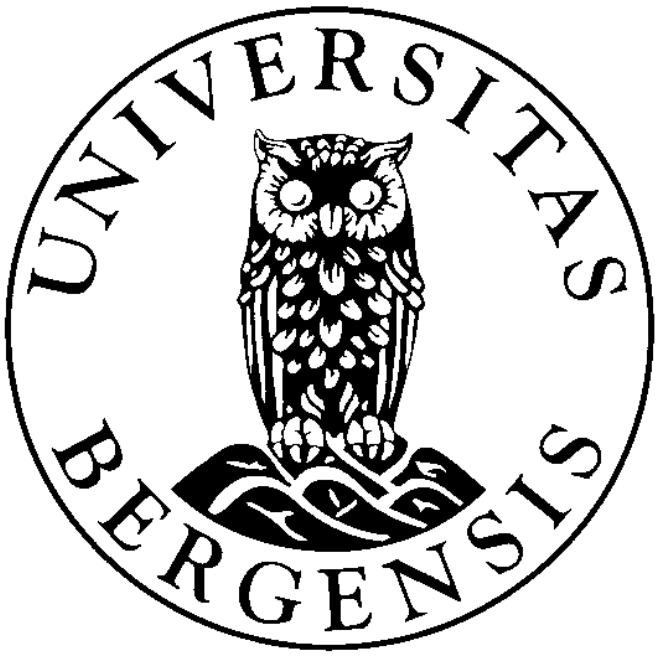
Priming adult beginner learners

A study of cross-linguistic lexical priming in German and Spanish learners of Norwegian



Master thesis in linguistics at the University of Bergen

Julie Matilde Stenhammer Sverreson

May 18th 2016

Acknowledgements

First of all, I would like to express my very great appreciation to my supervisor, Christer Johansson, for his advice and guidance throughout the process, and for the idea for the project.

I would like to offer my thanks to the University Library of Bergen, especially Ellen Solvik for the help I got with references and EndNote, and *Akademisk Skrivesenter* for proofreading the thesis. I would also like to express my appreciation to LLE (the department of linguistic, literary and aesthetic studies at the University of Bergen) for the financial support.

My special thanks are extended to all of my 48 participants for taking time to travel to the University and participating in my experiment. Without you, there truly would not be a thesis to write.

Victoria Troland, Eli Rugaard, Randi Brendbekken and Guri Kristin Rosendal also deserve my sincere gratitude for reading and commenting with their different perspectives on different parts of the thesis. Your help was invaluable. I also wish to acknowledge the language help, provided by Jerick Thompson Clarke, Klaas Erik Römer and Eli Rugaard, with translations of instructions and examples, and with making the German word pair list.

My parents deserve a thanks as well, not only for their support and encouragement, but also for their help with different parts of the process. My mother has shown an enthusiasm for my subject that has at times been greater than my own, and pushed me through hard times with a stream of more or less relevant research articles. My father helped with organization and planning and has been a great help with handling data in Excel.

I would also like to thank my fiancé, Jakob, for the patience and support, encouragement and comfort.

A big thanks to all my friends for being there when I needed them and understanding my need to not always be there when I needed that. A special thanks to all six of you in *Felleschatten* for being who you are.

Last, but not least, I would like to express my deep gratitude to Christina Lisa Stremme for paving the way.

Abstract

In a previous experiment, Stremme, Troland, and Johansson (2014) found that beginner learners of Norwegian also experience an effect of meaning related words in a primed lexical decision task. The study tested the effect of Norwegian prime words on English L2 speakers living in Norway, but originating from different countries. The results showed a strong, significant meaning priming effect for beginner learners (Stremme et al., 2014).

The present study is a replication of this pilot study, controlling for two different L1 (rather than a common L2) and for target language proficiency. Target language proficiency was investigated through a vocabulary test. German and Spanish native speakers living in Norway were tested in a cross-linguistic primed lexical decision task. The test was designed to measure the priming effect of Norwegian prime words on L1 target words. An unprimed baseline was added to remove lexical effects (Stremme, 2015). The aim of the study was to investigate the connections between L1 and L2 in the mental lexicon of beginner learners, and examine how L1 might influence L2 learning.

The hypotheses of the study are (1) Norwegian words will have a priming effect on L1 meaning related target words for beginner learners of Norwegian and (2) German L1 speakers will experience a stronger priming effect from Norwegian prime words than Spanish L1 speakers, as German is more closely related to Norwegian than Spanish is.

The results show a significant priming effect of meaning related words (-19 ms) for the German L1 speakers [$F(1,54.3)=6.15$, $p<0.02$*]. However, no such effect was detected for the Spanish L1 speakers. This lends support to the hypotheses. The vocabulary test scores did not correlate with the priming effects, indicating a difference between implicit and explicit language knowledge, but for the German L1 speakers, a positive correlation between the reaction times and the vocabulary test scores was observed.

The results indicate that (1) the German L1 beginner learners had already made a connection between the meanings of the words in their L1 and L2 in their mental lexicon, and that (2) the German L1 speakers have an advantage over Spanish L1 speakers in learning Norwegian. Further research is needed to investigate the effect of L1 and other previous L2 in learning a new language as an adult.

Sammendrag

I et tidligere eksperiment fant Stremme et al. (2014) at også nybegynnere i norsk opplevde en effekt av betydningsrelaterte ord i en primet leksikalsk avgjørelsesoppgave (lexical decision task). Studien testet effekten av Norske primeord på Engeske andrespråksbrukere, som bodde i Norge, men kom fra forskjellige land. Rsultatene viste en sterk, signifikant betydningsprimingeffekt for nybegynnere (Stremme et al., 2014).

Denne studien er en replikasjon av den ovennevnte pilotstudien, som kontrollerer for to forskjellige førstespråk, i stedet for ett felles andrespråk, og nivå i målspråket. Nivået i målspråket ble undersøkt i en orforrådstest. Tyske og spanske morsmålsbrukere ble testet i en primet krysslingvistisk leksikalsk avgjørelses-oppgave. Oppgaven var utarbeidet for å måle primingeffektene av norske primeord på L1 målord. En uprimet baseline ble lagt til for å fjerne de leksikalske effektene (Stremme, 2015). Målet med studien var å undersøke forbindelsene mellom førstespråk og andrespråk i det mentale leksikon, og studere hvordan førstespråket kan påvirke andrespråkslæringen.

Hypotesene i studien er (1) at norske ord vil ha en primingeffekt på betydningsrelaterte målord for nybegynnere i norsk, og (2) at tyske morsmålsbrukere vil oppleve en sterkere effekt enn spanske morsmålsbrukere, ettersom tysk er nærmere beslektet norsk enn spansk er.

Resultatene viser en signifikant primingeffekt av betydningsrelaterte ord (-19 ms) for de tyske morsmålsbrukerne[$F(1,54.3)=6.15$, $p<0.02$*]. Ingen slike effekter ble funnet for de spanske morsmålsbrukerne. Resultatene fra orforrådstesten korrelerte ikke med primingeffektene. Dette peker mot en forskjell i implisitt og eksplisitt språkkunnskap. For de tyske morsmålsbrukerne ble det observert en positiv korrelasjon mellom reaksjonstider og resultater i orforrådstesten.

Disse resultatene indikerer (1) at de tyske nybegynnerne allerede hadde knyttet sammen betydningene av ordene i førstespråket og andrespråket i deres mentale leksikon, og (2) at de tyske morsmålsbrukerne har en fordel overfor de spanske morsmålsbrukerne i å lære norsk. Videre forskning er nødvendig for å undersøke effekten av førstespråk og andre tidligere lærte andrespråk på læring av et nytt språk som voksen.

Table of contents

Illustrations

Figures

Tables

Glossing abbreviations

3	third person
ART	article
DEF	definite
GEN	genitive
M	masculine
POSS	possessive
SG	singular

A	adjective
Dem	demonstrative
G	genitive construction
N	noun
Num	numeral
Rel	relative clause

Adv	adverbial
Neg	negation
O	object
S	subject
V	verb

Chapter 1

Introduction

It is commonly noted that adult German native speakers learning Norwegian as a second language progress quickly in their classes, compared to students with other native languages. Two interesting questions within the study of bilingualism are whether second language learning is different to first language learning, and whether the native language will influence the progress of the acquisition of a given second language.

In 2013, as part of the university course *Experimental psycholinguistics*, which is included in the master program in linguistics at the University of Bergen, a cross-linguistic (Norwegian to English) primed lexical decision task was performed on beginner L2 (or L3) learners of Norwegian (Stremme et al., 2014). The participants were exchange students at the university, originating from several different countries, all speaking English as a first or second language. The results showed a strongly significant advantage for meaning related word pairs when compared to unrelated word pairs, for the beginner learners. No such effect was found in form related word pairs (Stremme et al., 2014). The study is presented in detail in section 1.1.

The study described in this thesis is a replication of this pilot study, controlling for the participants' native language, by using participants with the same native languages (German and Spanish), and offline knowledge of Norwegian with a vocabulary test. In addition, the present study will include an unprimed baseline to better control for lexical effects, like word frequency, so that priming effects can be studied without interference (Stremme, 2015). The aim of the study is to further investigate priming effects for beginner language learners and to examine whether participants with different native languages produces different results.

The following chapter will go into more detail on the background for the research questions (section 1.1), then present the hypotheses of the present study (section 1.2), before presenting definitions of some key terms (section 1.3). The final section (1.4) is an outline of the thesis.

1.1. Background

While several studies have shown that a child is able to acquire new words after only one exposure (Indefrey & Gullberg, 2010), it is generally assumed that few adults possess this ability. It is a common view that children are better at learning language than adults (DeKeyser, 2013), although the extent, cause and nature of this phenomenon remain topics of debate (see further Birdsong, 2006; DeKeyser, 2012; Hyltenstam & Abrahamsson, 2003).

In 1967, Eric Lenneberg formulated the *critical period hypothesis* (in Singleton & Ryan, 2004). As Field (2011) states, a critical period refers to "a period in the life cycle when there is greater sensitivity to certain kinds of stimuli". In second language acquisition (SLA) research this concerns the idea that children are naturally more talented in acquiring new languages than adults. In the critical period debate, three important questions arise.

The critical period hypothesis states that acquiring a language automatically from mere exposure (not instruction) is only possible in a restricted period, named the *critical period*, lasting from age 2 until puberty (Lenneberg, 1967 in Li, 2013). This hypothesis was based on the belief that brain lateralization is completed at puberty, with the language function in the left hemisphere, and that not having reached this stage yet gave children an advantage in acquiring language. Not much was known about neural maturation when Lenneberg developed his hypothesis, and lateralization appears to take place at a much younger age than Lenneberg assumed (DeKeyser, 2013).

The existence, nature and progress of the critical period has been debated and researched thoroughly (see further Birdsong, 2006; DeKeyser, 2012; Monner, Vatz, Morini, Hwang, & DeKeyser, 2013)

More recent studies have found counter-evidence to the idea that adults do not acquire language like children do, and to the existence of a critical period. A study by Johnson and Newport (1989) indicated a gradual decline in the ability to acquire a second language.

Additionally, there are studies showing evidence that adults have the ability to quickly acquire language in a way that resembles children's language acquisition. Gullberg, Roberts, Dimroth, Veroude, and Indefrey (2010) conducted an investigation into how adults acquire their

first words in a new language. The participants were Dutch native speakers with no previous knowledge of Mandarin Chinese, which is typologically and lexically quite distant from Dutch.

The participants in Gullberg's et al. (2010) experiment were exposed to audio-visual input in the form of a weather report in Mandarin. The weather report lasted only for seven minutes. The participants were then tested in their segmental, phonotactic and lexical knowledge. It was found that adult learners exposed to only a limited amount of stimuli from a previously unknown language quickly picked up new words. Even without instructions "they are able to extract segmental, word-form related information, lexical meaning from the context and map it onto word forms identified[...]" (Gullberg et al., 2010). This evidence could point to adults using fast mapping to acquire new words similarly to the way children do.

In the pilot study introduced in the beginning of the chapter, Stremme et al. (2014) found other evidence of adult language acquisition occurring quickly. In a cross-linguistic (Norwegian to English) primed lexical decision task, beginner learners of Norwegian were shown to benefit equally from Norwegian meaning related primes as native Norwegian speakers. The effect of cognate facilitation is well-documented in lexical decision tasks (see van Hell & Dijkstra, 2002; Lemhöfer & Dijkstra, 2004). However, the effect is not previously documented in beginner L2/L3 learners.

The participants of the pilot study (Stremme et al., 2014) were 22 non-native Norwegian speakers living in Norway for a less than a year. English was their second, but not dominant language. Participants were given a self-evaluation form for their knowledge of English. All participants deemed themselves able to read an English newspaper, and reported that they used the English language every day. Proficiency in Norwegian was not tested in any way, but all participants had little or no knowledge of the Norwegian language. Nevertheless they exhibited a significant decrease in reaction times of 51 ($p<0.001$***) ms when presented with Norwegian prime words that were related in meaning to the English target words, compared to unrelated words (Stremme et al., 2014).

1.2. Research questions and hypotheses

The aim of the thesis is to further investigate the priming effect (see section 4.1.1) from L3 to L1, for beginner L3 learners of Norwegian (discovered in Stremme et al., 2014) and control

for L3 proficiency and the native language of participants. In the study, language is controlled for by having two groups of participants with the participants in the same group having the same native language – either Spanish or German. As the experiment controlled for native language, another change was made to the original experiment; target language was changed from the participants' L2 or strong L3 to L1. The new study also controlled for the participants' proficiency level of the target language, by giving them a vocabulary test.

The study was designed to provide answers to the following three questions. (1) Does a meaning or form related word in Norwegian prime its counterpart in L1 for recent learners? The priming effect from meaning related words, detected in the previous study, was quite strong (-51 ms), but the study did not control for participants' native languages or proficiency level in Norwegian (Stremme et al., 2014). The effects of form and meaning were tested through different word pair categories (see section 4.4).

(2) Will the participant's native language affect whether – or to what degree – there will be a decrease in reaction times when a target word is primed by a meaning related word? Controlling for language also provides an opportunity to examine potential differences between the two language groups. As German is more closely related to Norwegian, sharing large parts of the vocabulary and a great deal of the syntactic features (see section 3.1), the priming effects might be stronger for native German speakers than native Spanish speakers.

(3) Do higher vocabulary test scores correlate with a larger priming effect? In the previous study, the beginner learners were found to benefit equally as much from the Norwegian prime words as the native speakers (Stremme et al., 2014). This leads to the question of whether there is a relation between the proficiency level of a language and the priming effects obtained. The vocabulary test allows for an investigation into the relationship between implicit and explicit knowledge (see section 1.3) of Norwegian.

The hypotheses of the thesis are as follows:

H0: There is no significant priming effect of Norwegian words on L1 target words for neither German nor Spanish native speakers

H1:

a. Norwegian words will have a priming effect on L1 meaning related target words for beginner learners of Norwegian

b. German L1 speakers will experience a stronger priming effect from Norwegian meaning related prime words than Spanish L1 speakers

The theoretical backgrounds for the two hypotheses will be presented in chapter 2 (1a) and 3 (1b). If, as in Stremme (2015), a meaning relation leads to a positive priming effect it will be taken as evidence that the effect is achieved at a higher abstract level, such as the lemma level (see further section 2.1; Levelt, 2001).

If there is a difference in the priming effects depending on the participants' native language it will be taken as evidence that the participants' native language could affect how quickly in the learning process the mental lexicons of L1 and L2 are connected at this abstract level.

1.3. Terms and definitions

A person's native language (L1) refers to the language acquired implicitly as a child from family or caregivers (Field, 2011). In the case of children growing up in a bilingual environment, it is possible to have more than one L1. This is the definition adopted in this thesis. The notion of *dominant* language – the one that is preferred by the language user – has not been accounted for.

Second language (L2) is a term describing a language that is learned non-natively, at a later time than the L1 (Field, 2011). L2 is often used to encompass all languages acquired after the first, but sometimes it is necessary to specify whether a language is indeed the second or third (L3) or fourth (L4) language. In this thesis, Norwegian will be referred to as the participants' L3, although it might in reality be their L4 or L5. All participants spoke English in addition to their L1, but the majority of the participants spoke other languages as well. In a few cases, participants

had more than one L1. Especially among the Spanish speakers, there were participants having more than one L1, like Basque or Catalan.

Bilingual is a term used to refer to persons who are proficient in two languages (Field, 2011), but, similarly to the term L2, it is often used to refer to all people who speak any number of languages in addition to their native language. This thesis will adopt this definition, unless it is necessary for comprehension of the text to make a distinction between bilingualism and multilingualism. How proficient a person must be in their L2 to qualify as a bilingual is a subject of dispute. Luk and Bialystok (2013) propose that bilingualism consists of several dimensions, like proficiency and usage. The term bilingual will in the present work incorporate those people who are not equally proficient in both languages.

Language proficiency is a term used to indicate a person's general ability in a specific language (Field, 2011). Proficiency is a very wide term that includes anything from the ability to recognize words to speaking fluently. To narrow this term down, it can be divided into different types of language knowledge. For this thesis, it is important to note the distinctions between *implicit* and *explicit* knowledge and *declarative* and *procedural* knowledge. While the terms are often used interchangeably (explicit=declarative, implicit=procedural), Ellis (1993) distinguishes between them.

Explicit knowledge is the knowledge of which language users are conscious, while implicit knowledge is the intuitive knowledge that they possess. Declarative and procedural knowledge on the other hand concern the degree of control that language users have over their knowledge. Declarative knowledge is used with effort and control, while procedural knowledge is used automatically (Ellis, 1993). In the context of this study, the RT experiment will test implicit and procedural knowledge, while the vocabulary test will test explicit and declarative knowledge (see section 2.5).

In the Norwegian language there are two different written varieties, bokmål and nynorsk. The Norwegian words used in the experiment are all in bokmål, and all participants who had received classes in Norwegian had learnt bokmål. Consequently, in the present work, Norwegian will refer to bokmål.

In the lexical decision task, the prime and target word pairs are connected by different relationships. This thesis will consider different categories of word pairs defined by whether their

relationship is through form or meaning, both or neither. The word pair categories are cognates, false cognates, translations and unrelated words.

Carroll (1992) defines cognates as “lexical items from different languages which are identified by bilinguals as somehow being ’the same thing’”. In the present work, *cognates* will refer not to completely homographic translations, but translations with a significant form similarity. An example of this is the Spanish *papel* and the Norwegian *papir*, which share the first syllable, and both translate to *paper* in English. While Carroll’s definition does not exclude this interpretation, it is common to use the word ‘cognates’ to refer to completely homographic translations, but, as mentioned, this thesis will use the term in the wider sense, describe above.

In the same way Szubko-Sitarek (2014) defines false cognates as "interlingual homographs, but in this thesis, the term will be used to refer to nearly homographic non-translations, or *interlingual neighbours*. The false cognates are, like the cognates, related through form, but unlike the cognates, they share no meaning relation. Examples of false cognates are the German word *Bär*, meaning *bear* and the Norwegian *bær*, meaning *berry*.

Translations in this thesis will refer to word pairs that are related through meaning but not through form. Almost all words have translations in another language, and although cognates are also a type of translations, these have been excluded from the translations category in this study. Some words have alternative forms. In these cases it has been attempted to choose a neutral form. Words that have two alternative translations where one is a cognate, have been avoided if possible.

Unrelated word pairs are not similar in form or meaning. In this case, great care has been taken to ensure that the prime and target words indeed share no trace of meaning or form relation. See section 4.4, for more details on the word pair categories.

1.4. Thesis outline

The previous chapter has presented the research questions and the background for the choice of these research questions. It has also presented previous research on the topic and provided some definitions of key terms. The following section will outline the thesis.

Chapter 2 and 3 will present the theoretical background for the study. Chapter 2 presents different models of the bilingual and multilingual mental lexica. Chapter 3 concerns typology. It

describes the similarities and differences between the languages involved in the study and presents theories on how these similarities and differences might influence second language acquisition.

Chapter 4 describes the methodology of the study is thoroughly explained and Chapter 5 presents the results.

Chapter 6 will go into more detail concerning the results, and try to answer the research questions. The chapter will also contain interpretation of the results in light of previous research and the theoretical background presented in chapters 2 and 3.

Chapter 7 contains the conclusion of the thesis. A more detailed outline is included in the introducing paragraph of each chapter.

Chapter 2

The bilingual mental lexicon

All the words that we acquire through our lives are stored in our minds and organized in a *mental lexicon*, so that we can easily access them when needed. The mental lexicon is the speaker's mental store of vocabulary items (Field, 2011). This lexicon holds all the information we need to produce, perceive, and understand a word. That information includes a representation of *word form* and *word meaning* per word. The word form is its phonology or orthography, while the meaning is the concept of the word that is connected to the word form.

All of the information on individual words is stored in lexical entries. According to Levelt (1989), the inner structure of a lexical entry is organized in levels. Phonetic features combine to make phonemes, which again combine to make the different parts of syllables. A combination of phonemes makes a morpheme which in itself or combined with other morphemes will constitute a word.

Several models have been developed to explain how bilinguals organize two mental lexicons. Questions that arise in regard to the bilingual lexicon are whether the lexica are separate or connected, and whether a potential connection is between word form or word meaning. This chapter will discuss these questions and present some of the models of the bilingual mental lexicon.

Hypothesis 1a of this thesis proposes that "Norwegian words will have a priming effect on L1 meaning related target words for beginner learners of Norwegian". As explained in section 1.2, a priming effect from meaning related words will be taken as evidence for the two mental lexicons being connected at a conceptual level.

This chapter will go into detail on the theories underlying this hypothesis and present models for the bilingual and multilingual mental lexicon. Section 2.1 discusses how words are

stored in the mental lexicon. Section 2.2. and 2.3. presents different models of the bilingual mental lexicon, and section 2.4. presents a multilingual model of the mental lexicon. The final section, 2.5. describes different types of language knowledge.

2.1. Lexical access

When speaking we are dependent on quickly accessing the words in our mental lexicon, for what we are saying to make sense and be produced with fluency. Retrieving the words from our mental lexicon is called lexical access, and there are several models attempting to describe this process. This thesis will present Levelt's (2001) model of lexical access, as it provides the necessary concepts to understand the following models of the bilingual mental lexicon, and processability theory (see section 3.2.2).

Levelt (2001) developed a model representation of lexical access consisting of two main levels, *lexical selection* - the selection of the correct *lemma* - and *form encoding* - retrieving the correct articulation. A lemma is a mental representation of a word (Field, 2011). In other words, the lemma represents the concept of a word, not its phonology or orthography.

To explain the model Levelt (2001) uses an example of how a person, upon seeing a picture of a horse and being asked to name it, accesses the right word in the mental lexicon. He proposes two stages of the process. The first one is lexical selection, where the participant must choose between the possible words to describe the picture. The first part of this stage, perspective taking", is to focus on the right concept. Next the person has to choose the right lemma, i.e. the "corresponding lexical item in the speaker's mental lexicon" (Levelt, 2001). Although "horse" is the most obvious answer, it would also be possible to choose a more specific or general term, like "stallion" or "animal".

The next stage is what Levelt calls form encoding, which consists of different stages of accessing the right articulation of the selected lemma. This is the stage that accesses the form of the word. First, the right morphemic and phonetic codes are retrieved, and combined to make phonemes, which then combine to make the syllable structure of the word. At last the whole word form is accessible (Levelt, 2001).

Figure 2.1 shows Levelt's model of lexical access for speech production, but the model can also be applied in reverse; to locate a meaning (a lemma) when a particular form is seen or

heard (Field, 2011). In a lexical decision task, this is exactly the process that the participants must go through for each word. Upon seeing the letter strings on the screen, participants must first decode the form, to reach the concept.

The following presentations of models of the bilingual mental lexicon will assume the distinction between word representations (form) and concepts (meaning).

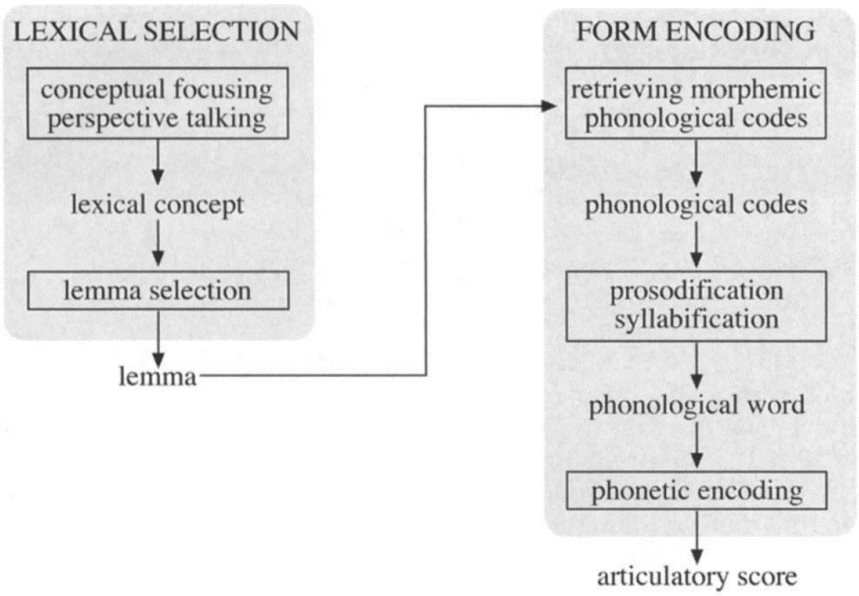


Figure 2.1: Levelt's model of lexical access (Levelt, 2001, p. 13465)

2.2. Shared and separate storage

So far this section has discussed how the mental lexicon is structured, but how does a second language fit in? There are several hypotheses concerning the structure of bilinguals' mental lexicon. These different hypotheses vary in how L1 and L2 are connected - if at all - in the mental lexicon.

Weinreich (1953) proposed three different kinds of organization of the bilingual mental lexica, the *coordinate, subordinate and compound* representations (see figure 2.2.). The representations are based on the Saussurean distinction between *signifier* and *signified*. The distinction is the same as the one mentioned in the introduction to the chapter; the signifier is the word representation or form and the signified is the word's concept or meaning. In the coordinate

(A) the two lexica are thought to exist separately from each other, with no contact between the words in L1 and L2. Each signifier has one signified. The compound (B), however is organized so that the signified of the two languages are treated as one. This model is similar to the concept mediation model (Potter, So, Eckardt, & Feldman, 1984), in which both L1 and L2 words are connected to the concept common to those words. In the subordinate (C), the signifier of the L2 word is stored as an extension of the sign in the L1. This representation implies a learning of L2 through L1. The model is similar to the word association model (Potter et al., 1984) which presents L2 words directly connected to their L1 equivalent.

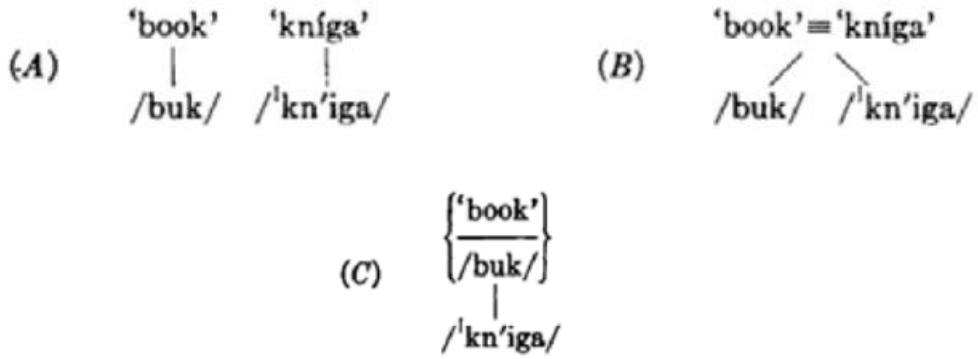


Figure 2.2: The three types of bilingual representation (Weinreich, 1953, pp. 9-10)

Kolers (1963) developed the *shared* and *separate storage models* based on Weinreich's (1953) representations. The separate storage model proposes that L1 and L2 are stored separately in the mental lexicon. This means there is a one-to-one relation between the meaning and the form of each word in each language, as in the coordinate representation (Weinreich, 1953).

Several different studies support this view. Bauvillain and Grainger (1987) studied context effects using a lexical decision task where the participants were to decide whether a word was English or French. They found that reaction times were slower if the word was presented in a mixed language list than if they were presented in a monolingual list. Participants took even longer in deciding if the word's orthography did not exclude one of the languages, and this was presented immediately after a word from the other language. The findings suggest that the participants were only able to access one language at a time, supporting the separate storage model.

Neurological studies have provided evidence for separate lexicons, showing how different areas of the brain are activated when using different languages (Perani et al., 1998). Observations of aphasia patients losing one language without any impairment to another also lends support to this view. (Albert & Obler, 1978).

Shared storage hypotheses on the other hand suggest that L1 and L2 are connected in the mental lexicon (Kolers, 1963). Shared storage does not specify how the languages are connected. Thus, this hypothesis is represented by both the compound and the subordinate models. This view is supported by studies of code-switching – the ability for a bilingual speaker to naturally jump from one language to the other within or between sentences – and linguistic interference, or *transfer* – L1 influence on L2 performance (see section 3.2.1; Singleton, 1999). This kind of interaction between languages is taken as evidence that the two languages of a bilingual have to be connected to each other.

Research on cognates and false cognates (Lemhöfer & Dijkstra, 2004) have also provided evidence supporting the shared storage hypothesis. Cognates are words from different languages that are similar in both meaning and form, and consequently present a unique opportunity to study the relationship between languages in the mental lexicon. False cognates are words which appear to be cognates, but do not share the same meaning.

Lemhöfer and Dijkstra (2004) performed four lexical decision tasks to compare reaction times for form and meaning related words. They found that shared orthography did not affect reaction times, unless it was combined with a shared meaning. Cognates, on the other hand were reacted to faster than the words in the control condition. The results support the conceptually mediated model.

A lexical decision task, as presented in this thesis (see section 4.1.1), provides a context for determining "whether only one or both languages are active when a string of letters is presented" (Kroll, Gerfen, & Dussias, 2010). The mental lexicon of a language must be activated to make a lexical decision in that language. In including two languages in the task, it is possible to investigate whether both lexicons can be activated at the same time.

Primed lexical decision tasks offer further insight into the organization of the bilingual mental lexicon.The priming effect is viewed as an activation of a mental representation of the prime word, spreading to the mental representation of the target word through a connection in

the mental lexicon (de Groot, 2013). Investigations into the priming effects of form and meaning similarities yield insight into the type of links that exist between the lexica.

2.3. The Revised Hierarchical Model

Combining the shared and separate hypotheses described above, there are several *mixed storage views*. The mixed storage views support both shared and separate storage at the same time. According to Levelt (2001) there are two separate levels of word representation - the lexical and the conceptual . In other words, there is one level for word forms and one for word meanings. In mixed storage views, the conceptual level is common to L1 and L2, but at the lexical level, the two languages stay separate (Szubko-Sitarek, 2014). This view could account for the findings in both Bauvillain and Grainger (1987), Perani et al. (1998) and Albert and Obler (1978) on one side, and Singleton (1999) and Lemhöfer and Dijkstra (2004) on the other.

One type of mixed storage view is Kroll and Stewart's (1994) *Revisd Hierarchical Model*. According to the Revised Hierarchical Model (RHM), each language is represented as a subsystem of one common language system. Thus, they are connected, but only one subsystem is activated at a time. In these subsystems there are conceptual links, connecting the words in each lexicon to conceptual nodes that represent that word's meaning (Kroll & Stewart, 1994).

In addition to how the two bilingual mental lexica are organized in relation to each other and across levels, it is also of interest to know how important external factors such as L2 proficiency and age of acquisition are to the organization.

Figure 2.3 is an illustration of the RHM. In this model L1 and L2 share one conceptual representation, similar to Weinreich's (1953) coordinate representation model, and in accordance with the concept mediation hypothesis (Potter et al., 1984). However, in the RHM, L1 is larger and shares a stronger link to the concept. The model combines *concept mediation* with *word association*, as there are lexical links between the L1 and L2 as well as the conceptual links. Kroll and Stewart (1994) argue that each of the models are active at different times in translation, depending on the direction of translation. They found that: *"*translation from the first language to the second is conceptually mediated, whereas translation from the second language to the first is lexically mediated" (Kroll & Stewart, 1994, p. 168).

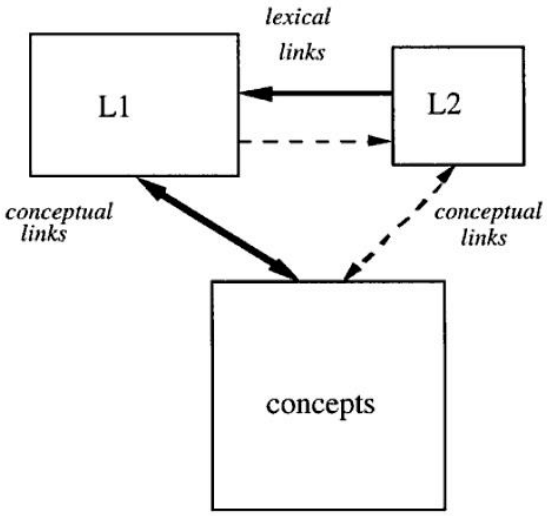


Figure 2.3: The revised hierarchical model (Kroll & Stewart, 1994, p. 158)

The argument is based on L1 to L2 translation being slower and less correct than L2 to L1 translation. Kroll and Stewart (1994) claim that translation from L2 to L1 operates at the lexical level and is thus faster. L1 to L2 translation will take longer because it is conceptually mediated. As figure 2.3 shows, accessing the conceptual level first is a detour.

A characteristic of the RHM is the asymmetrical structure. The model considers L2 proficiency and takes unbalanced bilinguals into account (Kroll & Stewart, 1994). According to the model, L1 and L2 are more strongly connected in the beginning of L2 learning. As the L2 proficiency increases, the conceptual links grow stronger.

The asymmetrical structure allows for a view of the bilingual mental lexicon as something dynamic rather than static, changing with the development and usage of the L2. This view is in accordance with the view of bilingualism itself being a continuum, as most bilinguals are not equally proficient in both languages (Grosjean, 2013; Luk & Bialystok, 2013).

Van Hell and Dijkstra (2002) support the importance of including proficiency level on the model. They argue that more frequently used words need less activation to be recognized, and, accordingly, a higher fluency in a language will affect activation in that language.

Brysbaert and Duyck (2010) have criticized the RHM as outdated, and point to new evidence. Among other things, they present evidence against separate lexicons and selective access (see Dijkstra, Timmermans, & Schriefers, 2000 for evidence from visual word

recognition). Kroll, van Hell, Tokowicz, and Green (2010), argue "that evidence for parallel access does not necessarily imply an integrated lexicon". In other words, the evidence that Brysbaert and Duyck (2010) point to does not exclude separate lexicons with parallel access and sublexical activation. For further insight into this discussion, see Brysbaert and Duyck (2010) and Kroll, van Hell, et al. (2010).

2.4. The multilingual lexicon

Bilingualism has been used as a term to describe a language users' knowledge of any number of languages. As explained in section 1.3, this thesis uses the term bilingualism in this wide sense. However, for the following section, it is necessary to make a distinction between bilingualism and multilingualism. The previous sections have presented several models for the organization of the bilingual lexicon. Although there has not been as much research on multilingualism specifically, the following section will present two models of the multilingual lexicon. Both models are based on Kroll and Stewart's (1994) RHM. Tymczyńska (2012) claims that the RHM is the only model of the bilingual mental lexicon that can accommodate a third language.

Gabryś-Barker (2005 in Szubko-Sitarek, 2014) developed a model of the multilingual mental lexicon, based on the RHM. The model proposes two types of links between the languages, conceptual and lexical links. Lexical links are word-to-word links, and conceptual links are links to concepts. Whether the words are accessed through the lexical or the conceptual links depends on a number of factors, such as language dominance, proficiency in each language, the task at hand, and the type of stimulus (Szubko-Sitarek, 2014).

As with the RHM, the multilingual model does consider proficiency. The model is illustrates the L1 lexicon as conceptually based with lexical links between all lexicons. With growing lexical competence, the lexical links grow weaker, but the conceptual links grow stronger. (Gabryś-Barker in Szubko-Sitarek, 2014).

Tymczyńska (2012) has also further developed the RHM in connection with an online translation experiment. She developed two models, one for professional interpreters, who have a high proficiency level in two or more languages, and one for her control group of non-

interpreters, who are more unbalanced trilinguals. For this thesis, the latter is more relevant, as the participants are beginner learners.

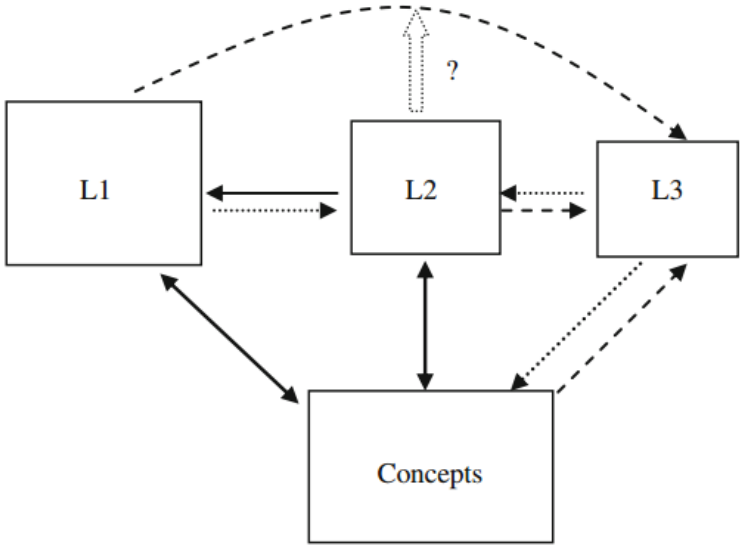


Figure 2.4: Model of trilingual mental lexicon (Tymczyńska, 2012, p. 157)

The model is in many ways similar to Gabryś-Barker's model. Conceptual links grow stronger with growing competence, as is illustrated by a weaker link from L3 to concepts. A difference between this model and Gabryś-Barker's model, is that the connection between L1 and L3 is illustrated as weaker than the other lexical links. The model thus reflects the sequence of language acquisition (Tymczyńska, 2012)

With regard to the present study, the links between the L1 and L3 lexica are investigated through a reaction time study (4.1). If a meaning priming effect is detected from L3 to L1, it would indicate that both lexica are activated at the same time (Kroll, Gerfen, et al., 2010). As the participants are beginner learners, links between lexica should, according to Kroll and Stewart (1994), be stronger than the L3 links to concepts. The RHM takes proficiency into account, with conceptual links growing stronger with increased proficiency (Kroll & Stewart, 1994). The participants' knowledge of Norwegian is tested through a vocabulary test. This is an offline test (see section 4.5.3), tapping into a different kind of language knowledge than the reaction time experiment.

2.5. Types of language knowledge

As described in section 1.3, there are several types of language knowledge. This thesis will be concerned with two major distinctions: explicit and implicit knowledge, and declarative and procedural knowledge. Ellis (1993) illustrates the interaction between these four types of knowledge in an L2 learning situation, as in table 2.1.

	Declarative	***Procedural***
Explicit	Type A: Conscious knowledge of L2 items	Type C: Conscious knowledge of learning, production and communication strategies. The learner can use the explicit knowledge easily and rapidly
Implicit	Type B: Intuitive knowledge of L2 items	Type D: Ability to employ learning, production and communication strategies automatically. The learner can use intuitive knowledge Fluently

Table 2.1: Implicit/explicit and declarative/procedural knowledge (Ellis, 1993, p. 94)

According to Ellis (1993), the starting point is usually type A knowledge in an L2 learning process. This knowledge comes from the information that is taught in the classroom. The goal is to reach type D knowledge, which means using the language both intuitively and fluently (Ellis, 1993).

Within the present study, these terms are used to describe the knowledge accessed in the two different tasks that the experiment comprises. Methods of testing implicit knowledge are called online methods (see section 4.1) and methods of testing explicit knowledge are called offline methods (see section 4.5.3). In a reaction time experiment, information is accessed quickly, automatically and intuitively. Thus, the task accesses implicit language knowledge in a procedural way (Type D).

The vocabulary test, on the other hand is an offline test of language knowledge. The test gives the participants time to think about their answers and make conscious decisions. The knowledge that is accessed through the vocabulary test is explicit and declarative (Type A).

Sonbul and Schmitt (2013) investigated the relationship between explicit and implicit lexical knowledge by having participants perform three different tests. Two of the tests were designed to test explicit knowledge; form recall and form recognition. A priming test was used to examine implicit knowledge. They found that there was a "clear disassociation between implicit and explicit lexical knowledge, and [that] one does not imply the other" (Sonbul & Schmitt, 2013, p. 151).

Chapter 3

Languages

As mentioned in chapter 1, the present study is designed to control for language, to further investigate the results of Stremme et al. (2014). The study hypothesizes about language that (b) "German L1 speakers will experience a stronger priming effect from Norwegian prime words than Spanish L1 speakers", because German is more similar to Norwegian than Spanish is.

This chapter will give an overview of the similarities and differences between Norwegian and German and Spanish respectively, and theories on how these similarities and differences can affect Norwegian L2 acquisition. The first part, 3.1. will present the languages in a typological context and describe their genetic classifications. In the following section, 3.2., theories of how these similarities and differences might affect the results of the priming experiment are discussed.

3.1. Language typology and genetic classification

The Germanic languages are split into two major groups; West Germanic and North Germanic (Hickey, 2012). German belongs to the West Germanic family and Norwegian to the North Germanic family. The two language groups are closely related, and share a great number of linguistic traits.

According to the World Atlas of Language Structures (WALS) online (Dryer & Haspelmath, 2013), German and Norwegian share much of the same phonological and morphological systems. Within nominal phrases the word order is identical in the two languages and negation is expressed in a similar manner. The syntax is also similar. Both languages are V2 languages.

Spanish belongs to the Romance language family, which consists of the languages deriving from Vulgar Latin (Encyclopædia Britannica, 2015). Comparing Spanish to Norwegian, the phonology shares fewer common traits. The same is true for nominal and verbal categories. Although the word order is approximately identical in normal declarative sentences (SVO), there are differences within nominal phrases and in other types of sentences.

In the following, some examples of the similarities and differences between the languages are presented. Lexica and syntax will be the primary focus of the comparison. Within phonology, the major differences are between Spanish and Norwegian, in the number of vowels and the syllable structure. Both systems are simpler in Spanish (Dryer & Haspelmath, 2013).

As mentioned, German and Norwegian are genetically closer than Spanish and Norwegian. This genetic closeness has resulted in the vocabularies of the two languages largely overlapping (Sandøy, 2000). The native Norwegian words are the words that stem from Proto-Norse. Proto-Norse in turn stems from Proto-Germanic (Store norske leksikon, 2009). As German also stems from Proto-Germanic, many of these native Norwegian words, will also exist in the German language. In addition to the native Norwegian words, the language also contains imported words.

According to Sandøy (2000), 44 percent of the words in the Norwegian dictionary are imported words. Low German was the third language on the list of exporting languages. 15 percent of the imported words came from Low German. The only languages that have exported a higher number of words to the Norwegian language are Latin (31 percent) and Greek (15 percent) (Sandøy, 2000).

Although the differences in syntax are not large between any of the three languages, German and Norwegian are closer than Spanish and Norwegian. The basic word order for normal declarative sentences is SVO in all three languages (Dryer & Haspelmath, 2013). Examples (1-3) demonstrate the same declarative sentence – *the man eats the apple* – in the three languages.

(1) Norwegian

mannen	spiser	eplet
S	V	O

(2) German

der Mann	isst	den Apfel
S	V	O

(3) Spanish
el hombre come la manzana
S V O

The structure of sentences (1-3) is the same, SVO. However, if an adverbial (A) – like *today* – were introduced, the structure of the Norwegian (4) and German (5) sentences would change to VSO, but the Spanish (6) sentence would remain the same (Dryer & Haspelmath, 2013), as shown in (4-6).

(4)
i dag spiser mannen eplet
Adv V S O

(5)
Heute isst der Mann den Apfel
Adv V S O

(6)
hoy el homre come la manzana
Adv S V O

All three languages express negation through a negative particle (Dryer & Haspelmath, 2013), but the placement of the particle in a sentence is not the same. In examples (7-9) the sentence from the above examples given in its negative form.

(7) Norwegian
mannen spiser ikke eplet
S V neg O

(8) German
Der Mann isst nicht den Apfel
S V neg O

(9) Spanish
el hombre no come la manzana
S neg V O

T*he prepositional noun modifier hierarchy* was developed by Hawkins (1983), as an implicational universal. The term implicational universal is used to describe a type of pattern

found in typological generalizations, where the universal can be structured in a hierarchy of implications, like this: 'If language X has structure Z, it will also have structure Y' (Croft, 2003).

In this context it will be used to demonstrate the similarities and differences between the three languages. The hierarchy reads like this: If a prepositional language places nouns before numerals (a), it will also place nouns before demonstratives (a), adjectives (b), genitive constructions (c) and relative clauses (d) (Hawkins, 1983).

All three languages are prepositional languages (Dryer & Haspelmath, 2013), and are thus subject to the prepositional noun modifier hierarchy (Hawkins, 1983). However the cutting point (a-d) is different for the three languages.

In the following examples (10-12), phrases which show these constructions in Norwegian, German and Spanish are presented, to demonstrate the similarities and differences in the word order of Norwegian, German and Spanish. While Norwegian (10) places demonstratives, adjectives and genitives before the nouns, German (11) places the genitive after the noun and Spanish (12) places adjectives and genitives after the noun.

(10)

a)	en	mann		
	a	man		
	Dem/Num	N		
b)	gammel	mann		
	old	man		
	A	N		
c)	mannens		hatt	
	man.DEF.M.GEN		hat	
	G		N	
d)	mannen	[som	er	gammel]
	man.DEF	who	is	old
	N	Rel.		

(11)

a)	ein	Mann
	a	man
	Dem/Num	N

b) alter Mann
old man
A N

c) der Hut des Mannes
ART hat ART man.GEN
N G

d) der Mann [der alt ist]
ART man [who old is]
N Rel

(12)

a) un hombre
a man
Dem/Num N

b) hombre viejo
man old
N A

c) el sombrero del hombre
ART hat POSS.ART man
N G

d) el hombre [que est -á viej -o]
ART man who is -3SG old –M
N Rel

In summary, Norwegian and German are, in addition to being genetically close, typologically more similar than Norwegian and Spanish. German also shares a large part of its lexicon with Norwegian.

3.2. Theories of L1 influence on SLA

The previous section has discussed some of the similarities and differences between Norwegian and the two target languages in this study. In the following, these similarities and differences will be viewed in relation to second language acquisition. Two different views on L1 influence on second language acquisition will be discussed. The first is the full transfer

hypothesis, and the second is the developmentally moderated transfer hypothesis, which is based on processability theory.

3.2.1. Transfer

Cross-linguistic influence is a term describing the different ways in which a language system may interact with another and affect the performance or development of the other language. Cross-linguistic influence includes, among other phenomena, transfer (Cenoz, Hufeisen, & Jessner, 2001). Transfer is the tendency of L2 learners to apply aspects of their L1 to their L2 (Field, 2011). Transfer can be both positive and negative, meaning that the aspects transferred might be either correct or incorrect in the target language.

Interlanguage is a term to describe the stages which a language learner goes through in their development in an L2 (Field, 2011). In other words it is all stages of language production between the beginning of the learning process and reaching native-like fluency. If transfer occurs, the interlanguage will contain structures from L1 and L2 at the same time.

Schwartz and Sprouse (1996) developed the Full Transfer/Full Access hypothesis, claiming that "the initial state of L2 acquisition is the final state of L1 acquisition" (Schwartz & Sprouse, 1996, pp. 40-41). In other words, the entire L1 grammar constitutes the base for the developing L2 grammar. In practice, this means that the L1 grammar is transferred to the interlanguage grammar until L2 grammatical structures are acquired and can replace the L1 structures (Schwartz & Sprouse, 1996).

The full transfer hypothesis predicts that both the German L1 speakers and the Spanish L1 speakers will transfer all of the L1 structures to their interlanguage grammar in the early stages of acquiring Norwegian. This transfer gives the German L1 speakers an advantage in that more of the transfer will be positive transfer, due to the similarities in the language.

Ard and Homburg (1992) suggest that positive transfer of lexical items could be of aid in acquiring words that are not similar. They propose that the L2 learners with a larger overlap between L1 and L2 vocabulary do not need to focus on acquiring the similar words, and have more capacity for the more difficult words than other learners of the same L2 with a more different L1.

Cenoz (2001) found that typology is a strong prediction of cross-linguistic influence. In a study on transfer comparing Spanish and Basque L1 speakers learning English as a L2, all of the participants presented a stronger influence from Spanish, which is typologically closer to English than Basque is.

Kim (1987) found that language transfer might affect language competence as well as language production. Additionally, there might be some typological universals that could predict the ease or difficulty with which a language structure is acquired.

(Falk & Bardel, 2011) have presented evidence that in L3 acquisition transfer is more likely to occur from L2 than from L1. They propose that the L2 functions as a 'filter' between L1 and L2 inhibiting transfer.

3.2.2. Processability theory

Processability theory (PT) was developed by Manfred Pienemann (1998) as a theory that can predict which structures that a language learner can process at a given level of development. The theory states that "structural options that may be formally possible, will be produced by the language learner only if the necessary processing procedures are available" (Pienemann, 1998, p. 4). In other words, a language learner can only process structures that they are already familiar with. According to the theory, language learners restructure their interlanguage step by step, gaining more and more conformity with L2 structures.

According to Pienemann (1998), the development of L2 acquisition is predicted by a universal implicational hierarchy of processing procedures. In other words, there is a fixed order for the acquisition of each structure. The order is as follows:

- Lexical access
- Category procedure
- Noun phrase procedure
- Verb phrase procedure
- Sentence procedure
- Subordinate clause

(Pienemann, Keßler, & Itani-Adams, 2011)

The first stage is lexical access, which entails identifying and acquiring the L2 words, and entering them into the mental lexicon. At this point, the language learner cannot process more than single words or fixed phrases. The next stage is assigning a category (noun, verb, agent, patient etc.). Canonical word order is used, as the language learner is still not able to process sentence structure. The third stage is phrasal procedure, divided into noun phrases and verb phrases. At this stage the language learner can process phrases, but the canonical word order is maintained. Next is sentence procedure, which implies acquiring grammatical functions and constructing main clauses with correct L2 word order. The last stage is processing subordinate clauses. (Pienemann, 1998).

The hierarchy is implicationally ordered, meaning that the processing procedures developed at one stage is a prerequisite for the procedures to be developed in the following stage. If a language learner is able to process sentence structure, he would also be capable of processing phrases and words. According to (Pienemann et al., 2011) the structure of the hierarchy is based on the stages in language generation. By this, they mean that a phrase cannot be produced without lemma access and a sentence cannot be structured without knowing phrases.

Once the theoretical framework of PT has been established, it can be applied to different L2s. Hypothesis Space is a term describing the constraint of PT on the grammar of a given L2 (Pienemann, 1998).

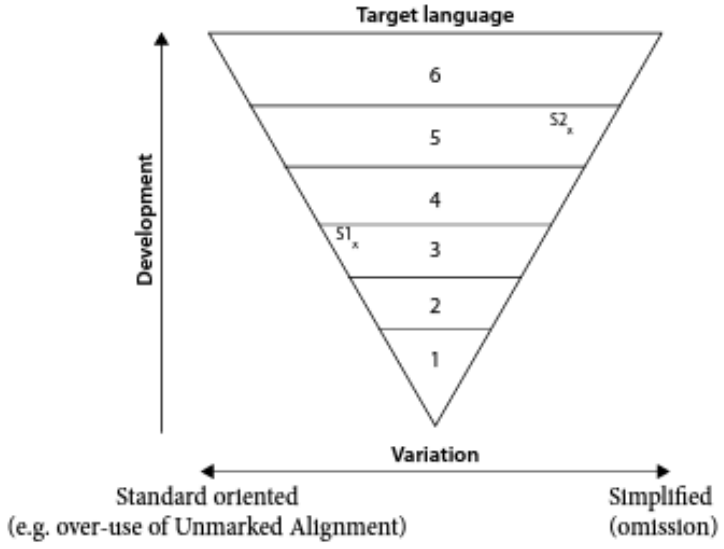


Figure 3.1: Hypothesis space (Pienemann & Keßler, 2011, p. 69)

Hypothesis space considers interlanguage variation. Variation is seen as one of two dimensions affecting the interlanguage grammar. The other dimension is development, which is expressed through the processability hierarchy, described above. (Pienemann & Keßler, 2011). Figure 3.1 demonstrates hypothesis space with each dimension on a separate axis. The stages in the processability hierarchy are presented from 1 to 6 on the axis of development, moving closer to the target language (L2) grammar with each stage. The variation axis displays different variational solutions that might be chosen by a learner. The interlanguage of a given L2 learner is marked as a point in this space (Pienemann & Keßler, 2011), depending on the learners' development and choices within the variational possibilities.

3.2.3. The developmentally moderated transfer hypothesis

As explained above, processability theory disregards the learner's L1. This is contrary to the full transfer/full access hypothesis of Schwartz and Sprouse (1996). However, with the incorporation of the developmentally moderated transfer hypothesis (Pienemann & Keßler, 2011), L1 transfer is still considered a possibility, constrained by the processability hierarchy. The developmentally moderated transfer hypothesis (DMTH) predicts, among other things, that:

1. If the L1 and the L2 contain the same structure which appears late, this structure will not be transferred at the initial state
2. However, this constellation does imply an advantage over learners whose L1 does not contain the structure in question: the structure will be acquired more effectively once it is processable.

(Pienemann & Keßler, 2011)

In other words, the hypothesis states that transfer of a linguistic structure is only possible if the learner has reached the developmental level required to process the structure in question.

Evidence for the DMTH is found in Håkansson, Pienemann, and Sayehli (2002). They tested prediction (1) by investigating whether Swedish L1 speakers in the initial stages of learning German transferred the V2 structure to German. Both Swedish and German are V2

languages, but according to the DMTH, the structure should not be transferred at such an early stage, because it is not yet available to the language learner (Pienemann & Keßler, 2011).

In accordance with the prediction (1), the Swedish language learners did not transfer the V2 structure to German L2, but rather produced sentences with canonical (SVO) word order (Håkansson et al., 2002).

Prediction (2) is supported by the research of Haberzettl (2005), on the acquisition of the German split verb form by Turkish and Russian children learning L2 German. Turkish has a preferred SOV word order, while Russian is an SVO language. In the German split verb sentences, the inflected auxiliary verb occurs in the regular verb position, while the meaning bearing, non-inflected verb is placed at the end of the sentence (SVOV). The Turkish children acquired this structure more quickly and efficiently than the Russian children (Haberzettl, 2005). Pienemann and Keßler (2011) suggest that this facilitation in favor of the Turkish children is due to transfer of the placement of the verb in the final position.

Bardel and Falk (2007) criticize the DMTH, saying that "If the [interlanguage grammar] has to wait for a positive transfer effect until it has reached a particular processability level, then transfer itself becomes superfluous" (Bardel & Falk, 2007, p. 465). They attribute the findings of Håkansson et al. to what they call the *L2 status factor* – that transfer to L3 is more likely to occur from L2 rather than L1 (Falk & Bardel, 2011)

Bardel and Falk (2007) argue that the L2 status factor is stronger, than typological closeness between L1 and L3. They studied learners of L3 Swedish or Dutch with different first and second languages, and the results indicate that syntactic transfer occurs from L2 rather than L1. They state that "in L3 acquisition, the L2 acts like a filter, making L1 inaccessible" (Bardel & Falk, 2007, p. 480).

The present study is concerned with lexical access. Lexical access is the first stage in the processability hierarchy. In other words, the developmentally moderated transfer hypothesis predicts that transfer can occur at this stage. As mentioned in section 3.1, the German vocabulary largely overlaps with the Norwegian. If transfer does occur in the way that Pienemann and Keßler (2011) propose, it would give the German L1 speaking participants an advantage over the Spanish L1 speakers. If no such effect is detected, it would be interesting to perform a study controlling for L2 to investigate a possible effect of the L2 status factor.

Chapter 4

Method

Psycholinguistics aims to discover and explain the acquisition, production and understanding of language. A number of different methods, such as fMRI, PET and EEG scanning, have been used to study the anatomical placements and brain activation patterns associated with the these processes (Kaiser, 2013). Another common method for gaining insight into how we perceive language, which requires less expensive equipment, is reaction time (RT) studies.

This study uses a cross-linguistic open primed lexical decision task with Norwegian prime words and L1 target words measuring the reaction times of German and Spanish native speakers living in Norway. The overall goal of this study is to examine and test bilinguals in order to enhance knowledge about how bilinguals' mental lexicons are organized. The aim is to examine the priming effect of Norwegian for beginner learners of Norwegian.

The null hypothesis of the study is, as mentioned in 1.2 , that neither form nor meaning relations is affecting reaction times for either language.

The following chapter will present the methodology of the study. Section 4.1. introduces reaction time studies, and the lexical decision task in particular. Section 4.2. presents the participants in the experiment. Section 4.3. describes the context for the performance of the experiment. In section 4.4., the stimuli is presented. Section 4.5. and 4.6. concern the design and the performance of the experiment. In section 4.7., the methodology for the data processing is described, before a final section on the validity of the study in section 4.8.

4.1. Reaction time studies

Many of the processes underlying language processing are very rapid – so rapid indeed that a person is not conscious of the process taking place. It all happens in the course of

milliseconds. On-line methods are able to “tap into real-time aspects of language processing” (Kaiser, 2013). Reaction time studies are one type of on-line methods. In reaction time studies, a slower reaction time is usually associated with a heavier processing load. In other words, the more time that is needed to complete a task, the more difficult this task is believed to be. There are a number of different methods within reaction time studies, including self-paced reading and eye tracking, but this thesis will focus on lexical decision tasks.

4.1.1. Lexical decision tasks

Lexical decision tasks are a type of psycholinguistic experiment measuring reaction times, in which participants have to decide as fast as possible whether a given string of letters is a word (Collins English Dictionary, 2012). Participants looking at a screen are presented with words and non-words, successively, and must quickly decide which of the two categories each word belongs to by pressing one of two buttons. One button represents "yes" or "real word" and the other "no" or "non-word". The function of the non-words is to give the participants a task to focus on, and to prevent them from developing a strategy of always responding “yes” (Kaiser, 2013).

In a version of these tasks, the participant is presented with a prime word prior to each of these target words (de Groot, 2013). This version is called a *primed* lexical decision task. The term priming was first introduced in 1951, within the field of psychology, by Karl Lashley, who primed behavioral responses. Priming is understood as an increase in sensitivity towards a certain stimulus due to prior experience (Bargh, 2014). In other words, one type of stimulus, the prime, prepares the participant for a possible upcoming stimulus, the target.

In the case of word-to-word lexical decision tasks, both of these stimuli are strings of letters. The priming effect is viewed as an activation of a mental representation of the prime word, spreading to the mental representation of the target word through a connection in the mental lexicon (de Groot, 2013).

The primed lexical decision task was first performed by Meyer, Schvaneveldt, and Grant (1971). In their experiment, 12 high school students were presented with an equal number of associated and non-associated word pairs, and as many word pairs consisting of one real word

and one non-word. Meyer et al. (1971)'s initial use of this method in reaction time studies was influential to the field of psycholinguistics and the method has been further developed and used.

In primed lexical decision tasks, the prime word may be presented openly or masked. Open priming is the method originally used by Meyer et al. (1971) and it is still in use. In open priming, the prime word is usually detectable and readable to the participants. An open priming sequence might look like this:

prime (100 ms)
TARGET (1000 ms)

In masked priming, a mask (usually a row of hash marks) precedes the prime. The target directly succeeds the prime and both the mask and the target are presented for longer intervals. This method makes the prime word virtually invisible to the participants, although about one in fifty report still being able to see the prime word clearly (Forster, 2015). A masked priming sequence might look like this:

(500 ms)
prime (60 ms)
TARGET (1000 ms)

Although several studies have found that masked priming could actually increase priming effects (Forster, 2015), there are a number of difficulties associated with this method. Stremme (2015) suggests that the masked and unmasked priming paradigms measure different stages of lexical access.

Masked priming is often associated with short SOA (30-60 ms), Stimulus-onset asynchrony (SOA) is the time between the presentation of stimulus 1 (prime) and stimulus 2 (TARGET) (Kroll, Gerfen, et al., 2010). In other words, it is the amount of time that the prime word is presented on the screen. Stremme (2015) suggests that the word meaning is not activated with a short SOA A lack of activation at the conceptual level would inhibit a priming effect for meaning related words. Another problem with short SOA is that the prime might be missed completely if a participant blinks or moves his eyes.

Cross-linguistic primed lexical decision tasks consist of prime words and target words in different languages. This type of test is usually performed on bilinguals to gain insight into how bilinguals' mental lexicons are organized.

The present work describes a cross-linguistic open primed lexical decision task with Norwegian prime words. The reaction times of 48 non-Norwegian speakers living in Norway were measured to examine the possible priming effect of Norwegian prime words on target words in each participant's native language.

As mentioned in section 2.2, the priming effects can yield insight into the type of connections that exist between two languages in the mental lexicon of a bilingual. Controlling for language, and comparing Spanish and German, it is possible to investigate whether the German participants may have an advantage of positive transfer from their L1, in the experiment.

The following chapter will describe the participants, methods, design and statistics of the study. The outline of the chapter is as follows:

Section 4.2 explains the recruitment of participants and criteria for participation. Section 4.3 describes the location and environment of the experiment. Section 4.4 lists the stimuli used in the study and presents the word selection. In section 4.5, the instrumentation of the study is explained and in section 4.6, the procedures of the testing itself are listed. In section 4.7, there is a description of how the data was processed and analyzed. The final section, 4.8, deals with data validity and ecological validity of the study.

4.2. Participants

The majority of the participants were recruited through Facebook groups like "Deutsche in Bergen", "Hispanoablantes de Bergen" or "NHH Exchange students Spring 2015". Some of the participants learned about the experiment through posters at several of the university buildings and student houses. A few participants were informed about the experiment by previous participants.

For their participation each participant was given a compensation of 50 NOK (about 5 €). The reasons for compensating the participants was that the experiment took 30 minutes in total and, as a particular software was required for performing the experiment, several of the

participants had to travel by bus to the experiment facilities. The duration and location on top of the strict criteria for participation could potentially have made it difficult to get a large enough sample.

Along with information about the compensation, the time and place of the experiment and contact information, the inclusion and exclusion criteria were given in the text used to recruit the participants. The criteria are presented below.

Exclusion criteria. The participants should not:

- speak Norwegian fluently
- suffer from dyslexia or any other reading difficulties

Inclusion criteria. The participants should:

- speak either German or Spanish as a native language
- have lived in Norway for less than two years
- know enough Norwegian to order a coffee or something similar

As mentioned above, the participants were either German or Spanish native speakers. Two languages were chosen to examine the difference between these two groups' results. It might have been interesting to compare one of these groups to participants speaking a non-Indo-European language, but for a project of this size with limited time and resources, it was not feasible to find that many participants in the Bergen area. German and Spanish speakers both compose large groups in Bergen, especially among exchange students.

Additionally, it has been remarked that Spanish and German native speakers progressing at different rates, although there is very little empirical work on these groups.

4.3. Context

The experiments were conducted at the Faculty of Humanities at the University of Bergen, in a well-lit, soundproof room. The room is small, approximately 5 square meters.

Considering the door was shut, the air quality might have declined during a day, in terms of higher temperature, and possibly lower oxygen contents in the indoor air. To compensate for such effects, there is a ventilator inside. The ventilator produced a low background noise, which could be averted by a switch closing the ventilator. Whether the ventilator was left open or shut, was decided based on the air quality in the room before each experiment was initiated.

The room is located within another room, which serves as a classroom and a study hall. During the experiment the participants were alone in the room while the experiment supervisor was sitting in the study hall outside. In most cases there were people not connected to the experiment sitting in the study hall. Considering that the room was soundproofed, no noise from the outside room should have reached the participants during the experiment.

4.4. Stimuli

- Priming: Baseline/Experiment
- Relation: Form/meaning
- Language: German/Spanish

As mentioned above, the experiment was designed to measure priming effects in reaction times. The experiment was thus split into a primed part and an unprimed part. The unprimed part of the experiment functioned as a baseline. The reaction times from the unprimed baseline were used to eliminate interference from lexical factors, such as word frequency and word length. Balota and Chumbley (1984) found that word frequency is particularly influential on the reaction times of lexical decision tasks. With the unprimed baseline, priming effects could be separated from lexical effects in the primed experiment. In other words, the baseline "reset the clocks", for each word in the experiment. This process is further explained in section 4.5.2.

A difference in priming effect due to the relation between the prime and target words was investigated with the null hypothesis of neither form nor meaning relations affecting the reaction times (see section 1.2). The word pairs were divided into four categories based on their relation. The words were related either through form (false cognates), meaning (translations), neither (unrelated) or both (cognates). This will be discussed further in section 4.4.1 .

There were two participant groups, one with German L1 speakers and one with Spanish L1 speakers. The two groups were presented with matched experiments with their L1 as the target language, so that the results could be treated individually and compared with each other.

Each of the two experiments consisted of a primed experiment and an unprimed baseline (see section 4.5.2). The baseline and the primed experiment contained the same set of target words. To cancel out any repetition effect, the participants were divided into two groups. One group was presented with the baseline before the primed experiment and the other group with the primed experiment first.

4.4.1. Word selection

80 word pairs (Norwegian to L1), where each word was of relatively high frequency, were selected for each experiment (see Appendix A and Appendix B). Additionally 80 non-words were created per language. Van Kesteren, Dijkstra, and De Smedt (2012) found that bilinguals use language-specific letters or bigrams to correctly categorize a word as part of one of their languages. The non-words were generated in accordance with the orthography of each language respectively, to make it more difficult for participants to separate them from the real words in their native language.

In other words, in the Spanish experiment there were 80 Spanish words paired with 80 Norwegian words and 80 Spanish-looking non-words paired with 80 Norwegian-looking non-words. Consonantly, the German experiment contained 80 German words paired with 80 other Norwegian words and 80 German-looking non-words paired with the same 80 Norwegian-looking non-words as in the Spanish experiment.

The prime words were presented in lower case letters and the target words in upper case letters. The reason for presenting them in different letter case is to avoid a form overlap effect. This is especially important for the word pair categories that are related through form.

The word pairs were divided into four categories – based on whether they were related by meaning or form – per language: *unrelated, false cognates, translations* and *cognates* (See section 1.3). Table 4.1. displays the relations between the word pairs in each of the four categories.

	Cognates	***Translations***	***False cognates***	***Unrelated***
Meaning	Yes	Yes	No	No
Form	Yes	No	Yes	No

Table 4.1: Word pair relations

The *unrelated* word pairs are related in neither meaning nor form. An example of an unrelated word pair taken from the experiment is the Norwegian-Spanish word pair *skap* and *FLOR*, which share little resemblance in orthography or phonology, and translate to *cupboard* and *flower* respectively. Finding words for the unrelated category might seem simple, but for a people, it is a difficult task. Although most words are probably unrelated to each other, it is easier to think of a word that is related in some way. “This indicated that translation pairs are to some extend closer to min[d,] which implies that such selected words will have faster reaction times even without priming” (Stremme, 2015, p. 74).

The *false cognates* are related in form only. If a participant does not know the meaning of the Norwegian word, it might appear to them to be a cognate. The Norwegian-German word pair *kopp* - translating to cup - and *KOPF* - translating to head - is an example of false cognates included in the experiment. The word pair *ost* - meaning cheese in Norwegian - and *OST* - meaning east in German - are complete homographs and could not be included. The reason for not including homographs is that the participants should not experience any confusion related to which language each word was in.

The *translations* category simply included what the name implies. These word pairs were translations which did not share a form similarity, like the Norwegian-Spanish word pair *kvinne* and *MUJER*, both translating to *woman* in English.

As mentioned in section 1.3 cognates, in this thesis, will refer to translations with a significant form similarity as opposed to homographic translations, for the same reason that homographs were excluded from the false cognate category. This was done for the same reason as with the false cognates. To give an example, *gift* and *GIFT*, both translating to *poison* could not be included. An example taken from the list of word pairs used in the experiment is the Norwegian-German *katt* and *KATZE*, both meaning *cat*.

As described in section 3.1, the vocabularies of German and Norwegian largely overlap. As a result of this, it proved difficult to find word pairs that were not cognates, or near-cognates. As a result, the *translation* category includes some near cognates, such as *tommel* and *DAUMEN* and the *false cognate* category includes some words that are less frequent.

In general, the words chosen for this experiment were common, high-frequency nouns from each language. However, the frequency was not formally controlled for in the selection process. Some of the words, particularly in the false cognates category, were less frequent than the others measured in Google search hits. For instance, the German word *KATZE* has over 32 million hits, while *IMPFSTOFF* only has 525.000. Word frequency and familiarity might affect how quickly the word is accessed in the mental lexicon (van Hell & Dijkstra, 2002). For this reason, an unprimed baseline was included The baseline contained the same target words as the primed experiment did, but none of the prime words. Stremme (2015) argues that including a baseline per participant will more effectively reduce the lexical effects of each word, so that priming effects alone can be investigated (see section 4.5.2).

There were strict inclusion criteria for the words pairs. To be included in the experiment a word pair had to meet these criteria, listed below. Different dictionaries were used to find word pairs that met the criteria. For the false cognate category, online lists of false friends were used as an aid to find possible word pairs.

Inclusion criteria Every word in the list must:

- be a noun
- be between 1 and 3 syllables, and 3 and 9 letters.
- be relatively frequent in the language
- appear once and only once in the list

The reason for the upper limit of nine letters per word is that the participant should have enough time to process the whole word, make a decision and press the correct button within one second.

The average word length per category was supposed to be approximately the same, so as to avoid word length to factor into the reaction times. This was done per language. The German

words tended to have slightly more letters than the Spanish ones, due to differences in orthography in the two languages. On average, the number of letters in each word (including Norwegian words) in the Spanish experiment was 4.66, while in the German experiment, it was 4.73. In table 4.2 Let the average number of letters per category is presented.

	Unrelated	***False cognates***	***Translations***	***Cognates***
Spanish	4,7	4,55	4,65	4,725
German	4,725	4,325	5	4,875

Table 4.2: Length of words per category

4.5. Experiment design

4.5.1. RT experiment

The participants were tested in a cross-linguistic open prime lexical decision task. The prime words were Norwegian and the target words in the participant's L1.

The experiment was programmed in SuperLab 4.5, which is an experiment builder made to design psychological experiments measuring reaction times. This allows for easy programming of presentation time, different text types, and randomization. Data presentation was also facilitated by this program.

Each word in the experiment was marked with the relevant information in Superlab. This information included which language the target word was in, whether the relation between prime and target words was through meaning or form and whether the word was a real word or a non-word. Additionally, every word pair received a number. This included a letter representing each language (g for German and s for Spanish), a letter saying it was a real word (w) and a number.

The Cedrus response pad RB-540 (see figure 4.1.) was used in the execution of the task. The five buttons on this pad leaves less room for confusion and error from the participants' side than a computer keyboard’s more than 100 keys would have (Cedrus, 2015). This helped secure fewer incorrect answers and less hesitation and thus the inclusion of more of the data in the study

and a better measure of reaction times. Another important advantage of this response pad is that it has a 2-3 ms reaction time resolution, compared to 20-35 ms on a normal keyboard (Cedrus, 2015). *Superlab* is conveniently designed to connect the input from the response box to the correct response in the expermiment.

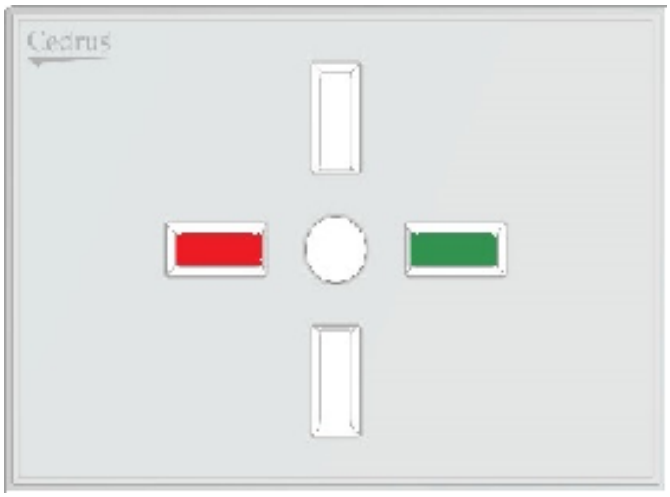


Figure 4.1: Cedrus response box RB-540

For this experiment, only two buttons were used, as the only possible answers were *yes* or *no*. These two buttons were coloured green for yes (real word) and red for no (non-word).

The presentation times were set to 100 ms for the prime words and 1000 ms for the target words. The prime word was programmed to stay on the screen for the full 100 ms while the target word would be presented for the 1000 ms or until the participant pressed a button on the response box. When a participant pressed a button on the response box, the experiment would move on the the next stimulus.

The prime words were written in lower case and the target words in upper case to avoid an effect of form overlap, especially in the cognates and false cognates categories. Both in the baseline and in the experiment each word (pair) was preceded by a * in the middle of the screen. This was used as a focus point for the participants to induce them to look directly at the word when it appeared. To avoid a detectable pattern in the word pairs, the presentation order was randomized.

4.5.2. Baseline

Stremme (2015) created a separate baseline for three priming experiments, to measure reaction times per target word, without presenting the prime words. This measurement can be used to investigate effects of word frequency and closeness to mind. The baseline RTs can be used to remove the lexical effects from the primed RTs, and thus allow for a more accurate measure of priming effects (Stremme, 2015).

In the present study an unprimed baseline was included per participant. The baseline contained the exact same target words as the primed part of the experiment, but none of the prime words. To avoid repetition effects from having been presented with the target words before, half of the participants were given the baseline before the primed experiment, and the other half were given the primed part first.

Originally, there were supposed to be four experiments only, two per language, one with the primed part following the baseline and one in the opposite order. During test rounds, however, it was discovered that the duration of the experiment (more than 30 minutes) was too long to complete without losing concentration, and the test subjects reported feeling tired and less attentive towards the end. This could have affected the reaction times of the last words being measured. Hence, each of the four the experiments was divided into two parts overlapping by 4 words per category.

In other words two experiments per language contained word pairs 1-12, from each word category – cognates, translations, false cognates and unrelated – while another four contained word pairs 9-20 from each category.

4.5.3. Vocabulary test

The final part of the test was a multiple choice vocabulary test containing all the 80 Norwegian words used in the relevant experiment. Stremme et al. (2014) did not control for target language proficiency in their experiment. Whereas the RT experiment is an on-line method of investigating the participants' implicit and procedural knowledge of language, the vocabulary test is an off-line method, investigating the participants' explicit and declarative knowledge of

Norwegian. Off-line methods in combination with on-line methods can provide crucial information about the final interpretations (Kaiser, 2013).

Like the experiment, the vocabulary test was programmed in *Superlab*. This time there was no time limit for how long the words could stay on the screen. The program would move to the next word only after receiving a response from the participant. The presentation order in this test was randomized. This was to avoid the detection of patterns in the relationships between the Norwegian words and their correct translation.

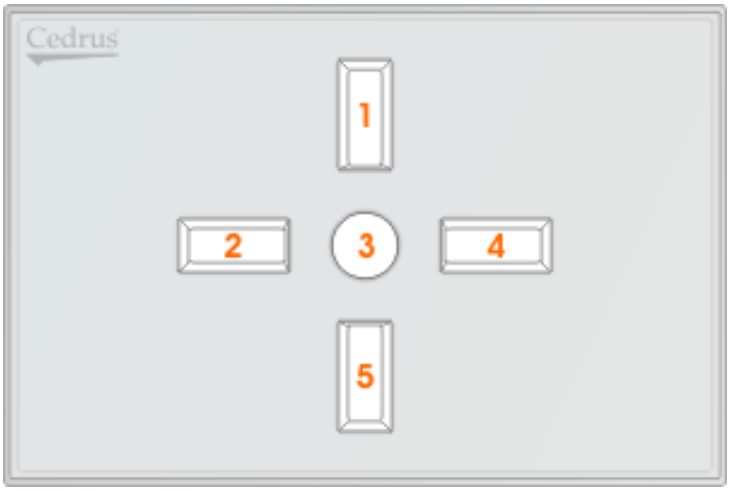


Figure 4.2: *Cedrus response box RB-540 for vocabulary test*

The Norwegian words were presented on the screen one by one, followed by a list of five words in the participant's native language, one of which was the translation. The prime word was always among the alternatives while the remaining alternatives were either false friends, completely unrelated or picked from the list of words used in the experiment. These words were selected to make it more difficult to guess the answer based on patterns in the test.

The response box was used in the execution of this task. In this case, all five buttons were used, as there were five alternatives for each word. The alternative translations were numbered and the buttons on the response pad were marked with corresponding numbers.

The vocabulary test was used as a cross-reference to the reaction times of each participant to see if there was a relation between a higher score and lower reaction times. It was also a possible way to assess inclusion. If a participant scored lower than random chance (20 per cent)

it would be understood as if he or she did not know sufficient Norwegian to have been primed by Norwegian words.

4.6. Performing the experiment

The test was divided into three parts; a questionnaire, the experiment and the vocabulary test. The total duration of the three parts was 30 minutes. Filling out the questionnaire and receiving instructions took approximately 5 minutes, the experiment itself lasted for 15 minutes and the vocabulary test was about 10 minutes, although the time spent on the vocabulary test varied slightly from participant to participant. The procedures will be explained in further detail below.

Firstly, each participant filled out a form of personal information including age, gender and native language (see Appendix C). Additionally they were asked to rank the languages they spoke by proficiency and give an indicator of their level of Norwegian. In addition to providing information about their native language, the participants were asked about their dialect in their native language. Both German and Spanish are spoken in several countries and consist of different varieties.

After filling out the form, the participants were given instructions on how to perform the task. They were told that they would be presented with words on the screen and that they had to decide if they were real words in their native language or not. They were informed that they had to be as quick as they could in their decision, and of which buttons they had to press. To avoid confusion, they were specifically asked not to press any button unless the word was in capital letters (target words). All participants were informed that they could leave the room at any moment, should they feel the need to do so. The instructions were given orally in English, but were repeated in the participant's native language in written form as part of the experiment. The repetition of the instructions was included both to ensure that the instructions would not be misunderstood, and to prepare the participants for reading in their native language. The participants were instructed to initiate the experiment when they were ready.

The participants were not informed of the aim of the study nor that what was measured was reaction times and not correct answers. This information was given after the participants had

completed the tasks. After receiving all of the instructions, the participants were left alone and the door to the room was shut.

Secondly followed the experiment - split into the baseline and the main experiment. The two parts of the experiment were presented with a break in between for the participants to rest their eyes so as not to lose concentration. The participants were in control of how long the break would last. They were instructed that whenever they felt ready to continue they should press any button on the response pad.

Finally, when the participants had completed the second part, the vocabulary test was initiated. Before the execution of the task, each participant was instructed both orally and in written form about the proceedings, including a demonstration of the new response pad. The ventilator was left open for the vocabulary tests, to avoid the air quality declining.

4.7. Data processing and statistics

4.7.1. Organization of data

As mentioned in section 4.7, Superlab effectively stored and organized the data. Each participant's results were saved with a participant number. This number included information about the participant's native language (G/S), whether the baseline or the experiment had been presented first (b1/e1), which part of the word pair list had been included (p1/p2) and single digit per participant in this group (1-6).

Once all of the data was collected, the files were easily converted to excel files. Excel facilitates the organization of the data, as separating it into different files and filtering out the irrelevant information, e.g. the reaction times of the non-words and incorrect responses.

All of the data was originally stored in one Excel file where each row represented one appearance of a target word, and each column contained relevant information about this particular word. The information included participant number, word number whether the word had been primed or not, language, meaning or form relation, word pair category, whether or not a word was repeated and reaction time. The data from two language groups were handled separately.

Before the data was analyzed, it was cut and reorganized extensively. Firstly, the non-words were removed. For this experiment the reaction times of the non-words were not relevant. Secondly, all words with reaction times lower than 350 ms were excluded. It is generally assumed that it is not possible to read and process a word, decide and press the correct button in less than 350 ms (Braun et al., 2006).

All incorrect answers were removed, as only the correct answers are relevant for the priming effects. The unprimed baseline and the primed experiment – were then divided into two distinct files to be handled separately. The two files were each split into two, so that the results from the two language groups could be analyzed separately.

The organized excel files, saved as CSV (Comma Separated Values) files, were accessed through *R*, a statistical computing and graphics software (R Development Core Team, 2008).

4.7.2. Baseline

As mentioned in 4.5.2, the function of the baseline was to remove the lexical effects from the experiment, to achieve a better measure of the priming effects (Stremme, 2015). Measuring RT per word in the experiment made it possible to take out the lexical effects, like word length, frequency and familiarity effects. The lexical effects were subtracted from the reaction times in the primed experiment, to 'reset the clock' for each word.

4.7.3. Experiment and vocabulary test

R is easily extensible by packages. For handling this data set the *lmerTest* package, which is based on lme4, was installed. Lmer stands for *linear mixed effect in R*, and the package allows for creating linear mixed-effects models. Mixed effects models offer several advantages to previous methods of RT analysis. According to Baayen and Milin (2010) mixed effects models are a method of data analysis which requires less a-priori data-trimming. The method allows for the consideration both variation in subjects (participants) and items (words) in one model (Baayen, Davidson, & Bates, 2008). The models were tested for significance using analysis of variance (ANOVA) (Winter, 2013).

Once the corrected values were calculated, the new *Experiment* file was uploaded to R. The lmer package was used to make a model to investigate the interaction effects between meaning and form per language.

The model estimates the effect of fixed and random variables on independent variables. The model includes fixed and random effects. Random effects are factors with levels sampled from a large population (Baayen & Milin, 2010). In other words, random effects can vary in a near infinite number of ways. Word number (Wnr) and participant are treaded as random-effect factor. Using word number and participant as random effects decreases the risk of underestimating variance.

Fixed effects are factors which have a limited number of levels (Winter, 2013). In the present study, form and meaning will be treated as fixed effects with two levels each, *yes* and *no*. This corresponds to the word pair categories, cognates (meaning-yes, form-yes), translations (meaning-yes, form-no), false cognates (meaning-no, for-yes) and unrelated (meaning-no, form-no).

In the linear mixed effects model, an intercept is assigned for each for each subject (participant) and item (word). Within the fixed effects, the intercept is the mean RT for the unrelated category, and the priming effects are measured in how much the other mean reaction times differ from the intercept.

Correlation effects were calculated using Spearman's rank coefficient correlation test (Spearman's rho). This test was used to investigate correlation between the vocabulary test scores and the reaction times and priming effects. The vocabulary test scores are measured in percentage of correct answers, while RTs are measured in milliseconds (between 350 and 1000) and priming effects can be both positive and negative values. A Spearman's rho can detect correlations in non-linear relationships.

4.8. Validity

The main advantage of the lexical decision task is that it is easy to administer. Large amounts of data can easily be collected and analyzed using this method (Baayen, 2014). Other advantages are that the method does not require expensive equipment and it facilitates data analysis. However, there are issues with the ecological validity of this method. Kaiser (2013)

points to the fact that the task is different from natural language processing, in that lexical decisions are not part of that process.

Primed lexical decision tasks are still considered a good method of investigating the internal structures of the mental lexicon. A priming effect is considered as an indication of parallel activation of the prime and target words.

Most of the participants of the study were between 18 and 35 years old. Only seven were older, and only one of them was older than 45. The majority of the participants were exchange students in Bergen. In other words the group was quite homogenous. It is well known that a larger and more heterogeneous sample provides more precise sample measures (Buchstaller & Khattab, 2013). On this account, the group might not be representative for the entire German and Spanish speaking adult population. However, a homogenous group allows for controlling for many other possible differences between participants that may affect reaction time more than language processing.

Baayen and Milin (2010) have found that older participants, on average, are slower in lexical decision tasks than younger participants. Participants of approximately the same age could make the results easier to analyze. A participant with generally slower RTs, may have more visible priming effects, because they have more room for ‘improvement’ in their RTs.

Chapter 5

Results

Two matched experiments will be described in this chapter; one which was presented to the Spanish speaking group and one which was presented to the German speaking group. Both experiments were primed lexical decision tasks with the priming direction L3-L1, where L3 was Norwegian and the participants were all beginner learners. Reaction times were measured for 80 word pairs, divided into four categories based on their relation through either form or meaning. The categories are cognates, translations, false friends and unrelated. In both experiments each L1 word was presented twice, once with the L3 prime and once without. The unprimed part was used as a baseline and the two parts were presented in the opposite order for half of each language group.

The participants produced a total of 4968 correct answers out of 5187 possible (=96 percent). There were 112 incorrect answers and 107 time-outs producing no response.

The following chapter will present the results of the study. The first section (5.1) describes the participants. In section 5.2, the results of the Spanish RT experiment are presented, while the German RT experiment results are presented in section 5.3. Section 5.4 deals with the results and implications of the vocabulary test.

5.1. Participant demographic

There were 48 participants in the study, 24 Spanish L1 speakers, and 24 German L1 speakers. Of them, 19 were men and 29 were women. They were aged 18 to 60. A demographic of the sample is given in table 5.1.

	Total	***Spanish***	***German***
Participants	48	24	24
Gender			
Male	19	10	9
Female	29	14	15
Age			
18-25	29	11	18
26-35	12	8	4
36-45	6	5	1
46-55	0	0	0
Over 55	1	0	1

Table 5.1: Participant demographic

All of the participants in the study were non-Norwegians who had lived in Norway for less than two years. Most of the participants had lived in Bergen for the duration of their stay in Norway, and the 44 who had taken classes in Norwegian had learned *bokmål*. Four participants reported that they had never received classes in Norwegian. Four participants reported speaking a variety of their native language that differs greatly from the standard, two Chilean Spanish speakers and two Swiss German speakers.

The participants spoke between one and four languages in addition to their L1 and Norwegian. All participants reported speaking English as one of their second languages.

5.2. Spanish RT experiment

Twenty-four Spanish L1 speakers participated in the experiment. Fifteen originated from Spain, three from Mexico, two from Venezuela, two from Colombia and two from Chile. Twenty-two had attended classes in Norwegian, from one month to one year. Five participants reported not having noticed the Norwegian words at all. Among these five participants, two had received less than two months of classes.

5.2.1. Main model

The reaction time experiment was a cross-linguistic lexical decision task. It consisted of eight identically designed experiments, within two groups, one with Spanish target words and one with German target words.

The Spanish experiment included 80 prime-target word pairs , where the prime word was Norwegian and the target words were Spanish, and 80 prime-target non-word pairs constructed in accordance with Norwegian and Spanish orthography. The word pairs were divided into four categories of 20 words based on their relationship; translations, cognates, false cognates and unrelated words. For each language group, there were four experiments. Each participant was presented with only one experiment. Each target was presented twice; once in an unprimed baseline and once in the primed experiment. Two of the four experiments presented the unprimed baseline before the experiment and the other two experiments presented the experiment first. Each participant was presented with 48 of the 80 word pairs and 48 of the non-word pairs. Two of the experiment parts contained word pairs 1-12 from each category and the other two contained word pairs 8-20 from each category.

The results of the four experiment parts were analyzed together in R (R Development Core Team, 2008) using linear mixed effects models (see section 4.7.3). Only correct answers of a word-word pair were considered in the analysis of reaction times. The 24 participants produced a total of 2692 (97 percent) correct answers. There were 45 incorrect answers and 48 no responses. The mean reaction time for correct responses was 599 milliseconds.

In the first group of the Spanish L1 speakers, presented with the baseline first, and supposed to contain the first part of the word pairs, all 80 word pairs were included. This was unplanned for. To investigate whether this group were affected by the longer duration of their experiment, this group was compared to the other groups. There was a slightly lower percentage of correct answers among group 1. Group 1 had 95 percent correct responses while the three other groups had a total of 98 percent correct responses. The mean RT for the correct responses for group one was 587 ms. The other three groups were on average slower, taking 603 ms on average to respond. In other words, while group one had fewer correct responses, reaction times did not increase due to the longer duration of the experiment.

The priming effects were calculated using a linear mixed effects model in R. Word number (Wnr) and participant were used as random effects, and meaning and form were the fixed effects. Interaction effects between meaning and form were also investigated The intercept is the mean RT for the unrelated word pair category, as the unrelated word pairs are neither connected through meaning nor form.

*>model_s <- lmer (rt ~ Meaning*Form + ((Meaning*Form)|Participant) + (1|Wnr))*

The effect of the meaning relation for Spanish speakers of -4.4 ms from the intercept (601 ms) was not significant ($p<0.68$).

5.2.2. Baseline correction

As mentioned in section 4.5.2, the Baseline was added to the experiment to remove the lexical effects from the reaction times, in order to study the priming effects only (Stremme, 2015). The unprimed baseline contained the same L1 words as the primed experiment and each participant was presented with both parts. The mean reaction time per word in the baseline was compared to the mean of every reaction time and the difference was subtracted from the reaction times from the primed experiment. After the lexical effects were corrected for, a new set of reaction times remained, in which any difference should be caused by external factors, such as word length or frequency was removed. The same model was applied to the new set of reaction times.

For the Spanish-speaking participants, there was only a very small effect of the meaning relation, lowering the reaction times by 2.2 ms from the intercept (604 ms), and it was not significant, [$F(1,35.3)=0.25$, $p<0.62$] .The effect of form relation was larger (-11.4 ms), although still not reaching significance, [$F(1,69.65)=0.3$, $p<0.59$].

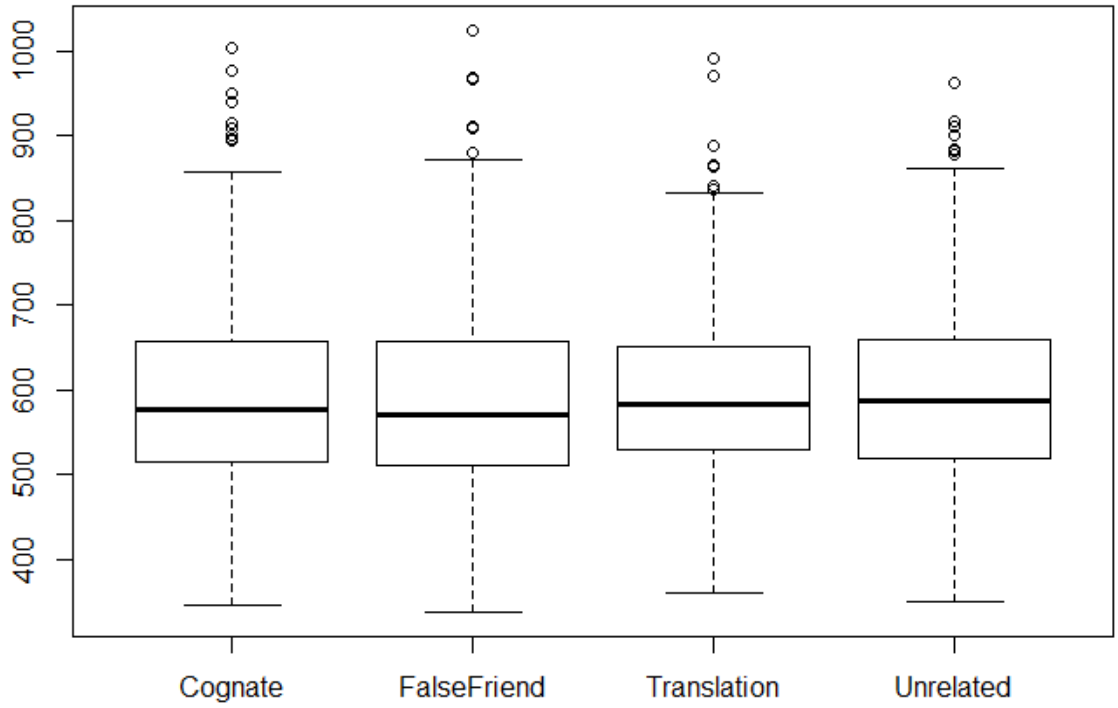


Figure 5.1: Boxplot of reaction times across categories, Spanish

5.2.1. Model accuracy

A normal probability plot was generated to compare the residual data – the difference between the observed data of the dependent variable and the fitted values – with the normal distribution. Figure 5.2 shows the residuals from the Spanish model as a line of connected dots, and the normal distribution as a line. The closer the residuals are to normal distribution, the more of the variation in the experiment, can be explained by the model. The model accounts for the results accurately up to the fourth quantile.

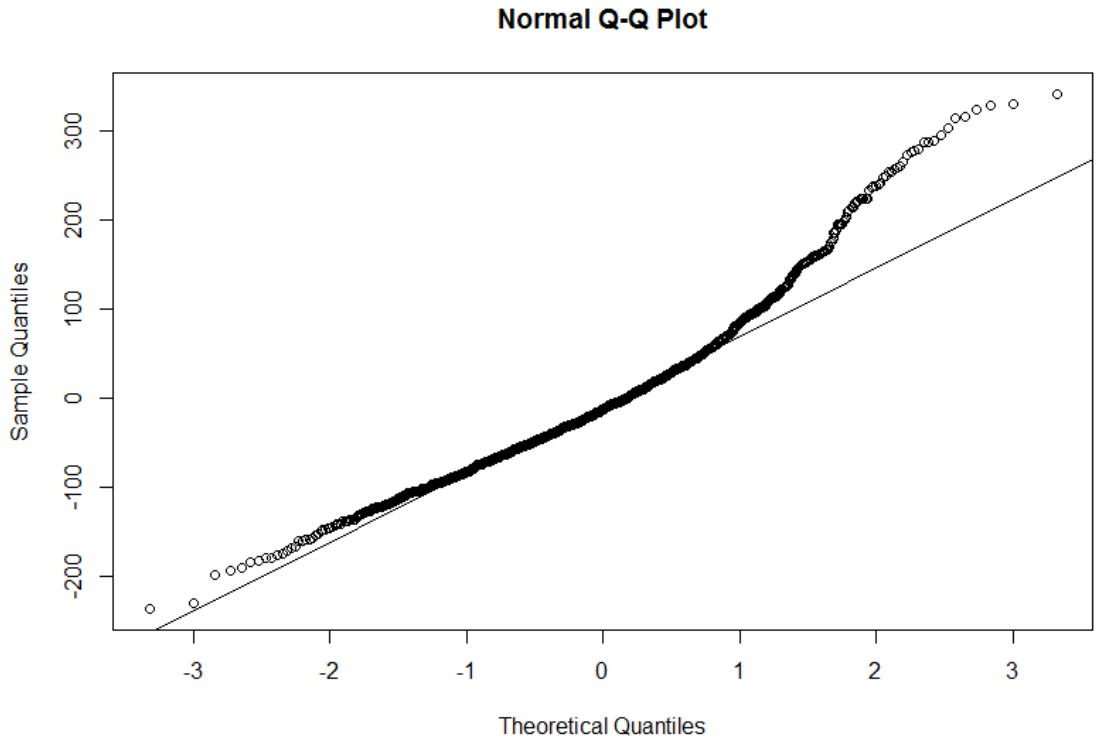


Figure 5.2.: Goodness of fit, Spanish model

5.2.2. Ventilator

The ventilator was left open during the RT experiment in some cases due to poor air quality in the test room. This caused a low, even background noise in the soundproof room. Among the Spanish speaking participants, 10 participants performed the experiment with the ventilator open, and 14 participants were given the test with the ventilator shut. The ventilator was always left open during the vocabulary test.

The effect of keeping the ventilator open or closed was investigated. There were no significant effects relating to priming effects. However, there was a significant effect on RTs for the Spanish L1 participants, which cannot be explained by the model. For further studies, it is advisable to control for the effect of the ventilator by keeping it in the same condition for all participants.

It is a possibility that the condition of the ventilator correlates with an unknown factor of the participants. Because of the small size of the group, any anomalous participant might have a large effect on the results.

5.2.3. Outliers

The mean reaction time of one Spanish speaking participant was more than 1.5 interquartile ranges (IQRs) above the third quartile when comparing the mean reaction times per participant per language.

After excluding the outlier, the effect of meaning increased to -5.8 ms from the intercept (600 ms), although the result is not significant, [$F(1,36.35)=0.38, p<0.85$]. The effect of form increased slightly (-12.4 ms), but the effect was not significant, [$F(1,70.77)=034, p<0.56$].

5.3. German RT experiment

There were 24 German L1 participants in the experiment. Among them, 21 originated from Germany, two from Switzerland and one from Italy. Two participants had not received classes in Norwegian. The other 22 had received classes for between two months and two years.

5.3.1. Main model

The German experiment included 80 prime-target word pairs, where the prime words were Norwegian and the target words were German, and 80 prime-target non-word pairs constructed in accordance with Norwegian and German orthography. The word pairs were divided into four categories of 20 words each, based on their relationship; translations, cognates, false cognates and unrelated words. The experiment was divided into four. Each target was presented twice; once in an unprimed baseline and once in the primed experiment. Two of the four experiment parts presented the unprimed baseline before the experiment and the other two presented the experiment first. Each participant was presented with 48 of the word pairs and 48

of the non-word pairs. Two of the experiment parts contained word pairs 1-12 from each category and the other two contained word pairs 8-20 from each category.

The results of the four experiment parts were analyzed together in R (R Development Core Team, 2008) using linear mixed effects models (see section 4.7.3)Only correct answers of a word-word pair were considered in the analysis of reaction times. The 24 participants produced a total of 2276 correct responses (95 percent). There were 67 incorrect responses and 59 no responses. The mean reaction time for correct answers was 598 milliseconds.

The priming effects were calculated using a linear mixed effects model in R. Word number (Wnr) and participant were used as random effects, and meaning and form were the fixed effects. The intercept is the mean RT for the unrelated word pair category, because the unrelated word pairs were neither related through meaning nor form.

*>model_g <- lmer (rt ~ Meaning*Form + ((Meaning*Form)|Participant) + (1|Wnr))*

The effect of the meaning relation for German speakers of -8,4 ms from the intercept (612 ms) was not significant ($p<0.66$).

5.3.2. Baseline correction

As mentioned in 4.5.2, the Baseline was added to the experiment to remove the lexical effects from the reaction times, in order to study the priming effects only (Stremme, 2015). The unprimed baseline contained the same L1 words as the primed experiment and each participant was presented with both parts. The mean reaction time per word in the baseline was compared to the mean of every reaction time and the difference was subtracted from the reaction times from the primed experiment. After the lexical effects were corrected for, a new set of reaction times remained, in which any difference should be caused by external factors, such as word length or frequency was removed. The same model was applied to the new set of reaction times.

For the German participants, meaning relation affected reaction times, [$F(1,54.3)=6.15$, $p<0.02$ *], lowering them by 19 ms from the intercept (617 ms). The effect of form similarity, [$F(1,52.6)=3.35$, $p<0.07$], of -12,7 ms, did not quite reach significance.

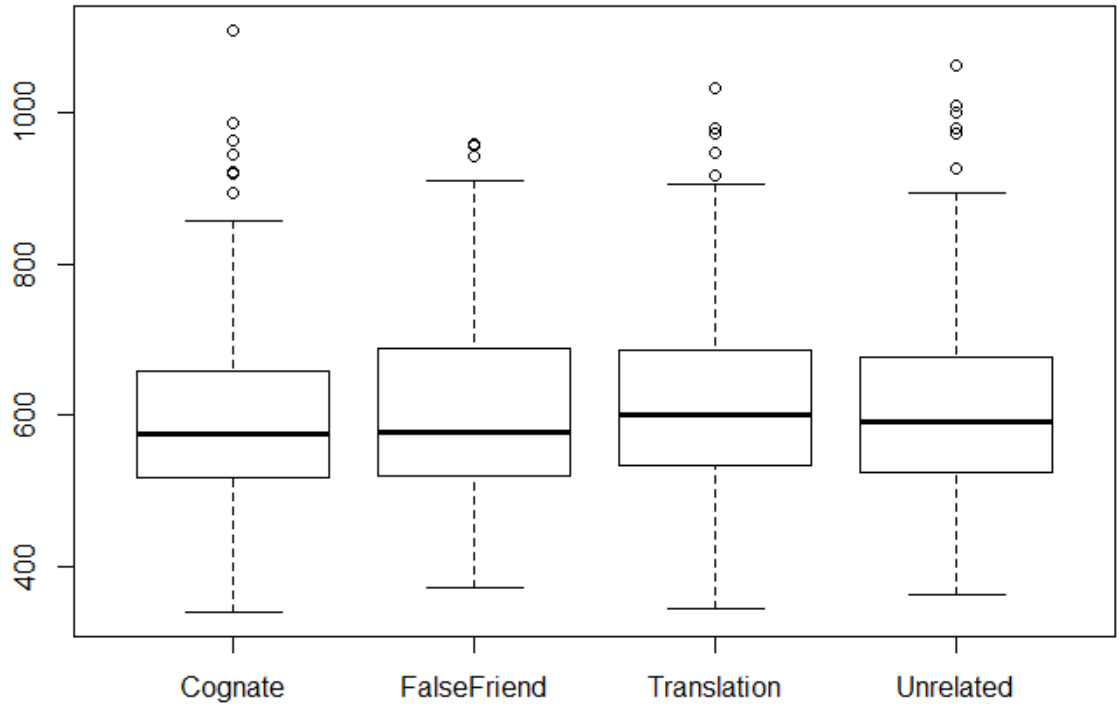


Figure 5.3: Boxplot of reaction times per category, German

5.3.3. Model accuracy

A normal probability plot was generated to compare the residual data – the difference between the observed data of the dependent variable and the fitted values – with the normal distribution. Figure 5.4 shows the residuals from the German model as a line of connected dots, and the normal distribution as a line. The closer the residuals are to normal distribution, the more of the variation in the experiment, can be explained by the model. The model accounts for the results accurately up to the fourth quantile.

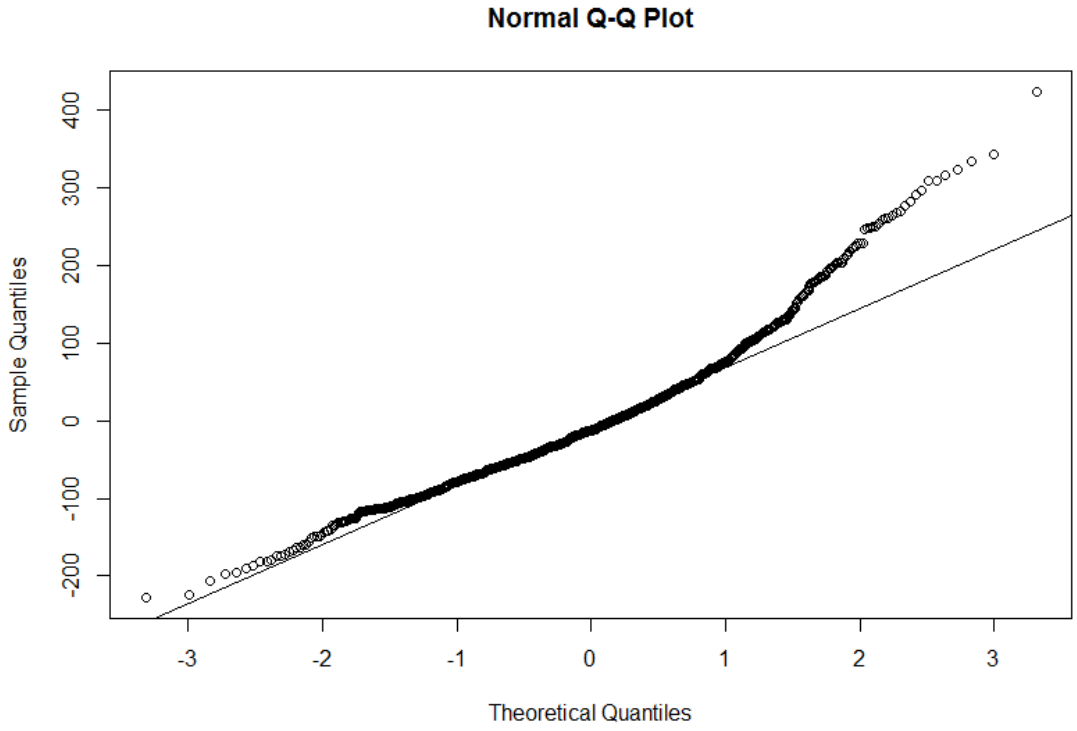


Figure 5.4: Goodness of fit, German model

5.3.4. Ventilator

The ventilator was left open during the RT experiment in some cases due to poor air quality in the test room. This caused a low, even background noise in the soundproof room. Among the German speaking participants, 9 participants performed the experiment with the ventilator open, and 15 participants were given the test with the ventilator shut. The ventilator was always left open during the vocabulary test.

The effect of the condition of the ventilator in the room was investigated. There were no significant effects of the ventilator on the priming effects for the German participants. This finding is equal to the what was observed for the Spanish L1 speakers. However, while there was a strongly significant effect on RTs of keeping the ventilator on for the Spanish participants, the the results show no effect on RTs for the German L1 speakers.

5.3.5. Outliers

The mean reaction time of one German speaking participant was more than 1.5 interquartile ranges (IQRs) above the third quartile when comparing the mean reaction times per participant per language.

When excluding the outlier from the German dataset, there was an increased effect of meaning, of -20.8 ms from the intercept (608 ms), which was significant, [$F(1,52.5)=6.54, p<0.01$*], and a slightly smaller effect of form (-12.5 ms) which was not significant, [$F(1,50.31)=2.86$, $p<0.10$].

5.3.6. Words

Among the German words there were six that significantly differed in reaction time from the mean reaction time. These words are presented in table 5.2 below. Word pair *gw65* and *gw33,* which are among these words, were the two word pairs in the German experiment with the fewest correct responses (50% and 58%).

Wnr	***Word pair***	***p-value***	
gw24	kone – EHEFRAU	0.02	*
gw33	vaksine – IMPFSTOFF	1.14e-05	***
gw4	atlet – ATHLET	0.0004	***
gw46	finger – PULLI	0.01	*
gw5	fetter – VETTER	0.0005	***
gw65	krutt – KRAUT	0.03	*

Table 5.2: German outlier words

The word pairs in table 5.2 are spread out across the four word pair categories. The word pairs, gw24 and gw33, belongs to the translation category, while gw4 and gw5 are cognates. In the unrelated category is gw46, and gw65 is a false cognate. To ensure a complete balance, only

one word per category was excluded in the outlier analysis. The word pairs gw24 and gw5 were therefore still included.

The target words of gw24 and gw33 are among the longest included in the experiment. IMPFSTOFF contains eight letters, but there are three nine-letter target words included in the experiment. EHEFRAU is three syllables long, and there is only one other three syllable word in the experiment. IMPFSTOFF is among the less frequent target words.

Excluding the outlier words produced no significant results.

5.4. Vocabulary test

The vocabulary test was included to test the participants' explicit knowledge of the 80 Norwegian words included in the reaction time experiment. There were two vocabulary tests, one per language group. The test was a multiple-choice test. Each Norwegian word was presented with five words in the participant's native language. Statistically, a score of 20 percent or less could be the result of guessing.

As shown in table 5.3. below, the two groups scored quite equally in the vocabulary tests. On average, the German speakers answered correctly on 72 percent of the questions, and the Spanish speakers answered correctly on 71 percent of the questions..

	Correct	***Error***
German	1402	518
Spanish	1361	559

Table 5.3: *Number of correct and incorrect responses in the vocabulary test*

To investigate whether the participants' answers were in some way biased by the placement of the buttons on the response box, the percentage of correct and incorrect responses per correct button was calculated. The percentages are listed in tables 5.4. (Spanish) and 5.5 (German).

	Correct	***Incorrect***
Button 1 (top)	64%	36%
Button 2 (left)	72%	28%
Button 3 (center)	78%	22%
Button 4 (right)	72%	28%
Button 5 (bottom)	70%	30%

Table 5.4: Correct and incorrect responses per button, Spanish

For the Spanish speaking participants the percentage of correct responses per button ranged from 64 percent to 78 percent. A chi-square test shows no significance for placement of the buttons on number of correct responses.

$\chi^2(DF=4)=1.42, p<0.84$

For the German speaking participants the number of correct responses ranged from 64 percent to 82 percent. In both cases there was one button which had a higher percentage of error clicks than the rest, but it was not the same button for each language group. A chi-square test of the percentages show no significance, indicating that the position of the button did not affect correct response.

$\chi^2(DF=4)=2.24, p<0.69$

	Correct	***Erroneous***
Button 1 (top)	73%	27%
Button 2 (left)	74%	26%
Button 3 (center)	72%	28%
Button 4 (right)	82%	18%
Button 5 (bottom)	64%	36%

Table 5.5: Correct and erroneous responses per button, German

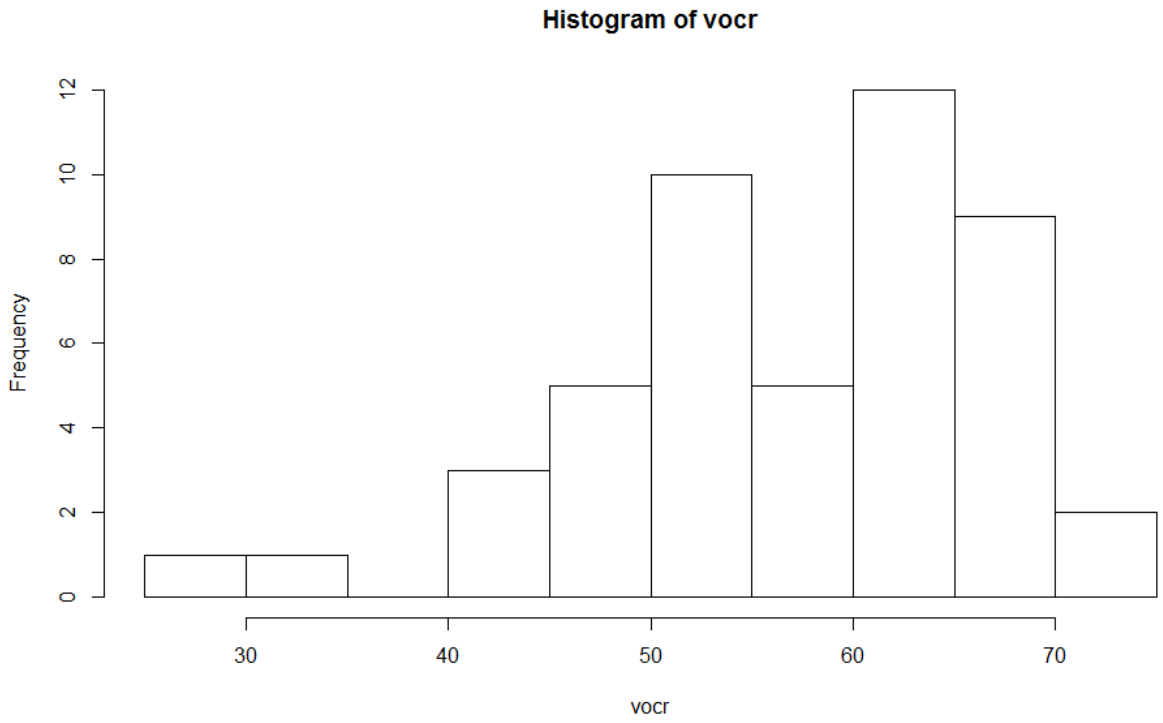


Figure 5.5: Vocabulary test correct responses (0-80) per participant

Figure 5.5: shows each participant's score. Two participants stand out, having less than 50 percent correct responses. One of the lowest scoring participants was German speaking and the other Spanish speaking. They were both among the four participants who had not received any classes in Norwegian. Neither of the two participants' results stand out in the reaction time experiment.

Although the results from the reaction time experiments were very different for the the two language groups, there were no large differences in their vocabulary test scores. As figures 5.6 and 5.7 below show, the distribution of the vocabulary test scores was similar for the Spanish and German groups. The figures divide the participants into four blocks depending on their performance on the vocabulary test. The blocks are 20-40, 40-60, 60-80 and 80-100 measured in the percentage of correct answers.

Among the Spanish L1 speakers were both the lowest scoring and the highest scoring participant. In other words, there was greater variation in the vocabulary test scores of the Spanish L1 speakers. Figure 5.6 Shows a near normal distribution of the test scores.

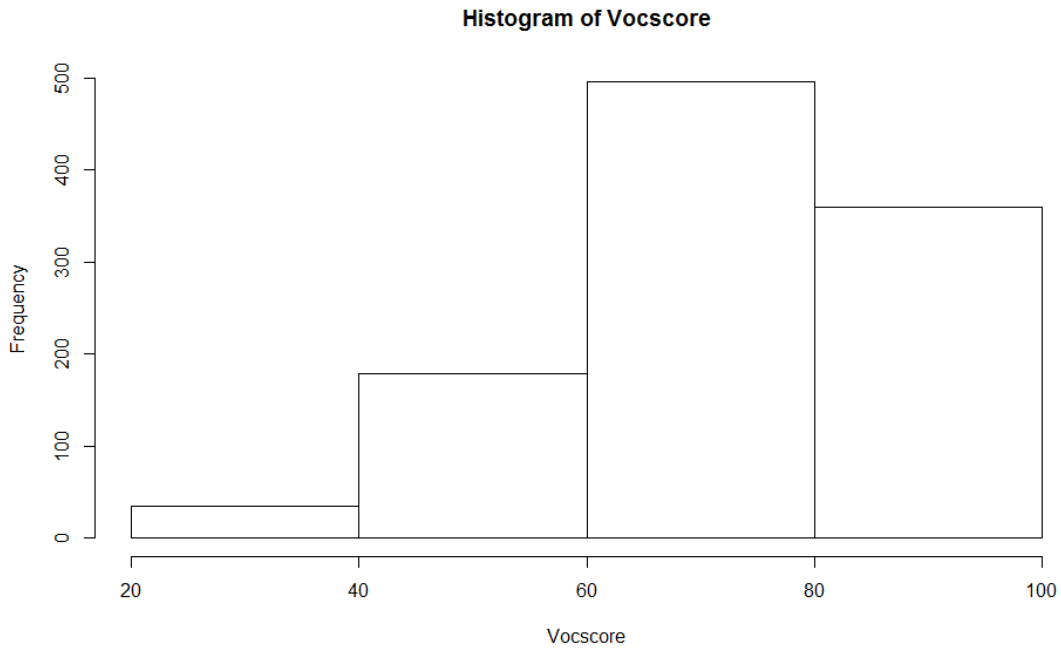


Figure 5.6: Spanish vocabulary test scores (in percent)

For the German L1 speakers, the distribution was similar, but with less variation. Figure 5.7 shows that a majority of the participants scored between 60 and 80 percent correct answers on the vocabulary test.

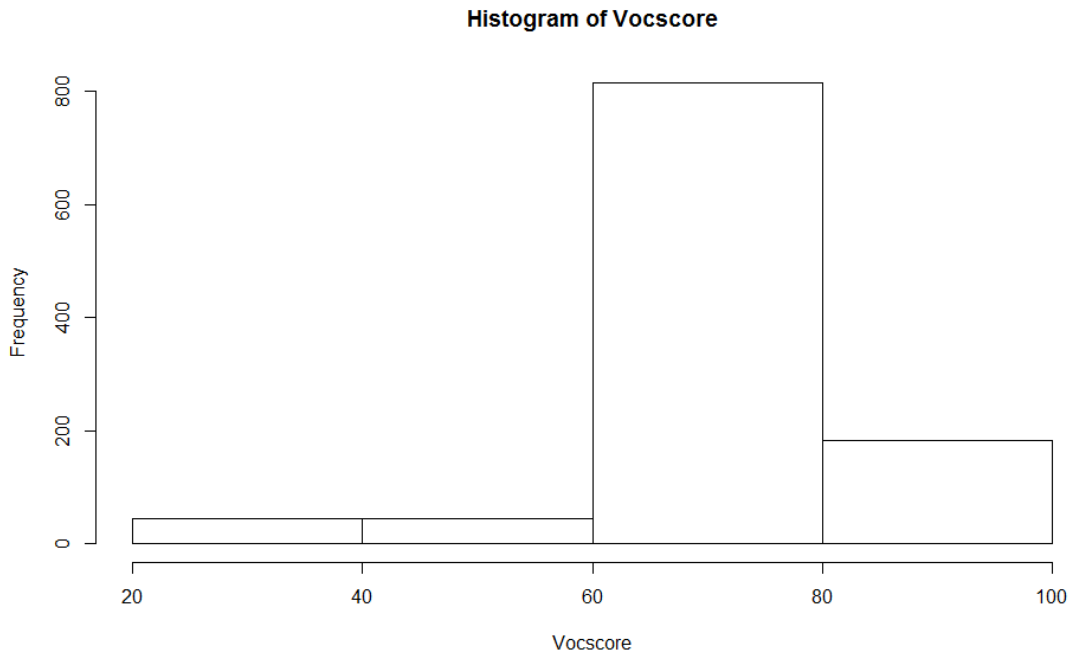


Figure 5.7: German vocabulary test scores (in percent)

Reaction times were measured for the vocabulary test as well. As with the test scores, the reaction times were similar for the German and Spanish groups. In general, the erroneous answers were responded to considerably slower than the correct answers. The fact that participants took more time to consider before making a wrong decision indicates that they hesitated more on these decisions. Table 5.6 shows the average reaction times in milliseconds for correct and erroneous answers per language.

	German	***Spanish***	***Total***
Correct	5740	6670	6198
Erroneous	11230	11131	11179

Table 5.6: Average reaction times in ms for correct and erroneous answers

The Spanish L1 speakers take around one second longer to answer correctly, than the German L1 speakers. With a total number of 80 question words, this adds up to an average of a little over one minute more per Spanish speaking participant to complete the test.

5.4.1. Correlation with reaction time experiment

As the data was not at the same scale, and could not assume a linear relationship, a Spearman's rank coefficient correlation test (Spearman's rho) was performed for correlation between the vocabulary scores and the results from the RT experiment. The vocabulary test scores are measured in percentage of correct answers, while RTs are measured in milliseconds (between 350 and 1000) and priming effects can be both positive and negative values. A Spearman's rho can detect correlations in non-linear relationships.

The results show a small decrease in reaction times ($r_s = -0.23$) for the German speaking high scorers in the vocabulary test, which was significant ($p<0.02$*). For the Spanish speakers, the correlation effect was not present ($r_s = 0.13$, $p<0.19$).

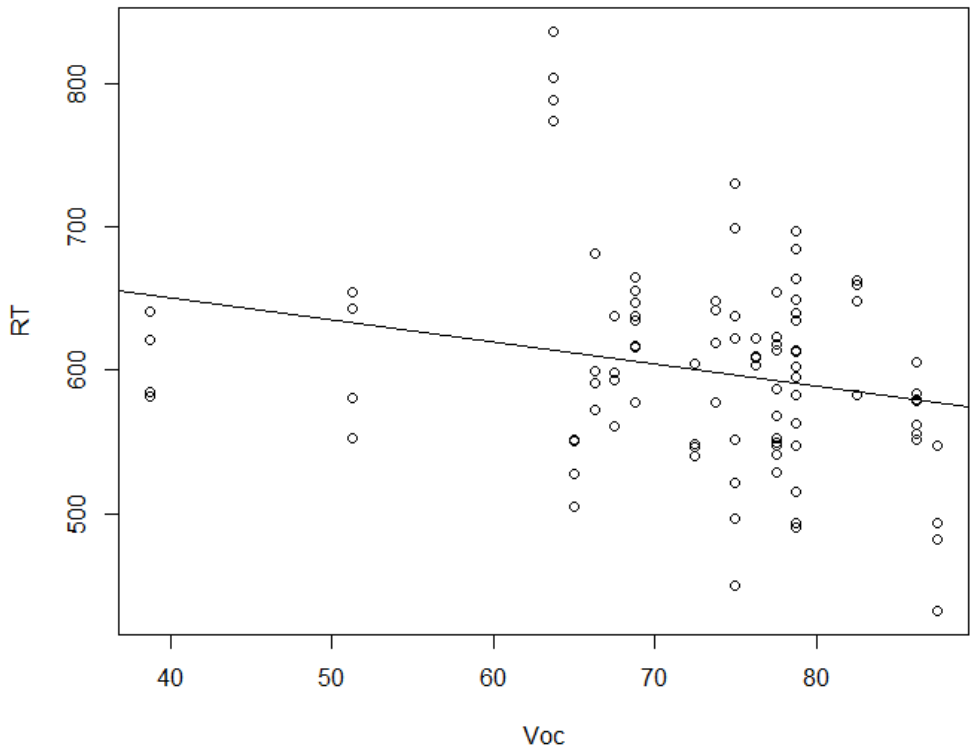


Figure 5.8: Correlation between RTs and vocabulary test scores, German

However, the priming effects are not predicted by the vocabulary test results. There was no significant correlation effect between priming effects in the lexical decision task and vocabulary test scores, in either direction, for the either the German or the Spanish speaking participants. In other words, the priming effects were not increased by the participants' explicit knowledge of Norwegian.

Figure 5.8 shows the correlation effect between RTs and the vocabulary test scores for the German L1 speakers. As the vocabulary test scores increase, there is a decrease in RT.

5.5. Summary

The results from the RT experiment show that German L1 speakers experience a small priming effect from meaning related Norwegian prime words even very early in the learning process. The finding is significant, although not strongly so. For the Spanish L1 speakers, there is no such priming effect for beginner learners.

Leaving the ventilator on during experiments significantly affected the Spanish participants' reaction times when compared to shutting it off, although the model does not explain this effect.

Six words in the German experiment stood out as significantly different in reaction time. Some of the words shared the trait of being among the longest in the experiment, either counting syllables or letters.

Table 5.7 Sums up the significant results from the RT experiments. There were no significant results for the Spanish group, but the German L1 speakers experienced a priming effect from meaning related Norwegian words. The effect and significance increase when the outlier participant is removed from the analysis.

	Intercept	***Priming effect***	***P-value***
German corrected RTs	617 ms	Meaning:-19 ms	P<0.02*
German without outliers	608 ms	Meaning: -20.8 ms	P<0.01*

Table 5.7: Significant findings in RT experiments

The vocabulary test scores were almost equally distributed between the two language groups. The two lowest scoring participants were among those who had not received any instruction in the Norwegian language.

The vocabulary test scores did not correlate with the priming effects from RT experiment. To summarize, the significant results of the correlation test were that while there were again no significant effects for the Spanish L1 speakers, there was a significant positive correlation between vocabulary test scores and RTs for the German L1 speakers (r_s = -0.23, ($p<0.02$*).

Chapter 6

Discussion

This study examined the relation between the first and second languages in the mental lexicon based on the priming effects of second language words on first language words. German and Spanish beginner L3 Norwegian speakers, living in Norway were tested in a cross-linguistic primed lexical decision task. The test was designed to measure the priming effect of Norwegian words on words in the participants' native language using an unprimed baseline. None of the participants had spent more than two years in Norway nor studied Norwegian for more than one year. On average the participants scored 72 percent on a vocabulary test of the Norwegian words used in the experiment.

The hypotheses of the study are (1) that a meaning relation between the prime and target words would produce a priming effect for beginner learners of Norwegian, and (2) that German native speakers would benefit more from this effect than Spanish native speakers due to the large syntactic and lexical overlap in German and Norwegian.

The results show a significant meaning priming effect for the German L1 speakers, but no significant effect for the Spanish L1 speakers. These findings support the hypothesis (1) for the German speakers, and to hypothesis (2). However, the Spanish L1 speakers did not only experience a smaller priming effect than the German L1 speakers. They experienced no priming effect at all. Tthere was no correlation between the scores of the vocabulary test and the priming effects.

These results indicate that (1) the German L1participantts had already made a connection between their L1 and L2 vocabularies through conceptual links and (2) that whether or how fast this connection is made depends on the language learner’s L1.

The following chapter will further discuss the results presented in the previous chapter in light of previous research and relevant theories on the subject. Section 6.1 will discuss the

research method. In section 6.2, the results from the RT experiment will be discussed, within the frame of the research hypotheses. Section 6.3 regards the vocabulary test scores in relation to the RT experiment. Finally, section 6.4 will discuss potential further research on the subject.

6.1. Experiment design

An important addition that was made to this study was the inclusion of a baseline per participant (see section 4.5.2).

In addition to the dependent and independent variables, there were some possible extraneous variables. Among these was the age of the participants. As mentioned in section 4.8, the RTs have been found to increase with increasing age of the participants in lexical decision tasks (Baayen & Milin, 2010). However, the majority of the participants were aged 18 to 35, with a few exceptions (see section 5.1).

Although the experiment controlled for native language of the participants, both Spanish and German are spoken in several different areas and consequently there are several varieties of the two languages. Four participants reported speaking varieties that differ greatly in lexicon from the standard varieties. Two of the German L1 speakers were Swiss and two of the Spanish L1 speakers were Chilean. If a participant spoke a variety with great lexical differences from the standard variety, this might affect the reaction times. However, none of the results indicate that these four participants differed from the other participants in RT or priming effects.

6.1.1. Ventilator

Two possible extraneous variables were the air quality and background noise in the room. The experiments took place in a sound proof room to avoid any disturbance. The room was small and, considering the door was shut during the experiments, the air quality declined during a day. Participants who participated at the end of the day might have experienced discomfort due to the poor air quality if the ventilator was shut. For this reason the ventilation was in some cases left on. This produced a small amount of background noise which might also have affected the reaction times (Kohfeld & Goedecke, 2013; Trimmel & Poelzl, 2006).

The effect of keeping the ventilator open or closed was investigated (see sections 5.2.2 and 5.3.4) to see if this in any way affected the results. There were no significant effects relating to priming effects. However, the reaction times for the Spanish L1 speakers were inexplicably affected by the ventilator.

Several researchers have investigated the effect of noise on different cognitive performance tasks. Research on background noise has come to different conclusions. Kjellberg and Landström (1994) point to evidence that the background noise caused by a ventilation system in an office environment causes workers to make more mistakes, and decreases learning rates.

Söderlund, Sikström, Loftesnes, and Sonuga-Barke (2010) on the other hand found that background white noise enhanced the performance of inattentive children in an episodic verbal free recall test. However, attentive children experienced a decrease in performance on the same test when subjected to background white noise. Wilkinson, Nicholls, Pattenden, Kilduff, and Milberg (2008) found evidence of decreases in reaction times in a visual memory test when participants were subjected to external noise.

The findings of the present study on the impact of background noise on reaction times and priming effects is inconclusive. Researchers on the field do not agree on the subject. For further studies, it is advisable to control for the effect of the ventilator by keeping it in the same condition for all participants.

It is a possibility that the condition of the ventilator correlates with an unknown factor of the participants. Because of the small size of the group, any anomalous participant might have a large effect on the results, if such a correlation exists.

6.2. RT experiment results

6.2.1. Hypothesis 1a

Hypothesis (1a) states that Norwegian words will have a priming effect on meaning related L1 target words for beginner learners. The following section will discuss this hypothesis in light of the findings in the present study, and in previous research. Some implications of the findings will also be discussed.

Stremme et al. (2014) tested 22 non-native speakers of Norwegian with English as a second language. The participants had different L1s and had lived in Norway for less than one year. The participants were given a cross-linguistic primed lexical decision task equal to the one used in this study, except priming went from L3 to L2, in stead of L3 to L1.

The results of the pilot study show a significant main effect of meaning [$F(1,20)=27.7;p< 0.001$***] and revealed a 51ms advantage for meaning related primes over the unrelated condition, 686ms. This effect was quite a lot larger than the one detected in the present study (Stremme et al., 2014).

When controlling for native language and L3 proficiency, in the present study, no priming effect was detected for the Spanish speaking participants (see section ??next for further discussion), and the effect of meaning priming was smaller (-19 ms) [$F(1,54.3)=6.15$, $p<0.02$ *] for the German speaking participants than for the participants in the first study. The Spanish speaking participants did not experience any effects at all.

The differences between the study of Stremme et al. (2014), and the present study are (1) that the present study controlled for the participants' native languages, (2) that the target languages in the present study were the participant's native languages instead of a strong L2 (English), and (3) that the present study included an integrated unprimed baseline per participant. The integrated baseline (3) was discussed in section 6.1.

Diminishing effects is a phenomenon that is often observed in many fields of research. Lehrer (2010) describes how the effects observed in a study gradually decrease with each replication study. The evidence is drawn from research on pharmaceuticals. As more tests are performed on a given drug, the effect of the drug seems smaller. Examples of the phenomenon from other fields of research, like psychology, are also described. One explanation for diminishing effects is to underestimate the variance in the population, by not using random effects where they should be used. Another explanation is control of more variables, as is done in the present study. Controlling for L1 (1) could be a possible explanation for the differences observed in the results of the two studies.

The reason that the target language was changed from L2 to L1,(2) was that the present study did control for the participants' L1. Controlling for L1 meant that using the L1 as a target language was possible, whereas in the pilot study it was necessary to encounter a common L2 (English) for the participants (Stremme et al., 2014).

However, there are studies which point to differences between the storage of L1 and L2. Brain imaging studies (Perani et al., 1998) have found evidence of L1 and L2 stimuli activating areas of the brain differently. In a PET investigation of unbalanced Italian-English bilinguals, it was found that Italian (L1) audio stimuli activated the temporal lobes and temporoparietal cortex more extensively than stimuli from English (L2). Thus, it is likely that there is a real difference in L1 and L2 in their associative networks. Although L3 was not investigated in the above-mentioned study, evidence of some separation between L1 and L2 was found. This evidence might suggest that all languages are separated, or it could indicate a separation between L1 on one side and all non-native languages on the other side.

Bardel and Falk (2007) found evidence of transfer occurring from L2 rather than from L1 to L3, even disregarding typological closeness between L1 and L3. They studied learners of L3 Swedish or Dutch with different first and second languages and the results indicates that syntactic transfer occurs from L2 rather than L1. They state that "in L3 acquisition, L2 acts like a filter, making L1 inaccessible" (Bardel & Falk, 2007).

If this is indeed so, then it might be that the L3 is more closely connected to the L2 than to the L1. This could explain the weaker results of the present study. Although the results from the German RT experiment indicate that L1 is not completely inaccessible from L3, it might be that it is more strongly connected to L2. This could lend support to a model of the multilingual mental lexicon based on the RHM (Kroll & Stewart, 1994) where all languages share a common conceptual base, but L3 is more closely connected to L2 than to L1, as in Tymczyńska (2012).

The present study does not control for participants' L2s or other L1s. They were asked to fill out any other languages that they speak in the questionnaire, and all participants spoke between one and four languages in addition to the prime and target languages and English was spoken as an L2 by all participants. However, they did not specify any proficiency level. The trilingual mental lexicon model (Tymczyńska, 2012) could be further complicated by the number of languages that a participant knows. Additionally, the argument that the L2 is more strongly connected to the L3 than the L1 is (Bardel & Falk, 2007) cannot be investigated properly without controlling for L2.

6.2.2. Hypothesis 1b

Hypothesis (1b) states that the German L1 speakers will experience a stronger priming effect form Norwegian meaning related prime words than Spanish L1 speakers. This was hypothesized because German is more closely related to Norwegian than Spanish is. The following section will discuss this hypothesis in light of the findings of the present study, and previous research. Implications of the findings will also be discussed.

The results show a small but significant priming effect for the German participants. It was hypothesized that the German participants might benefit more from the Norwegian prime words, but the fact that the Spanish L1 speakers experienced no priming effect at all, was surprising.

Processability theory (Pienemann, 1998) and the developmentally moderated transfer hypothesis (Håkansson et al., 2002; Pienemann & Keßler, 2011) predict that transfer can only occur when the structures in question are available to the language learner. According to the processability hierarchy, lemma access is the first stage in language learning. The DMTH thus predicts that transfer of lexical items would be a possibility.

The vocabularies of German and Norwegian largely overlap with cognates being more common than non-cognates (see section 3.1). According to the DMTH, German L1 speakers should have an advantage over the Spanish L1 speakers in acquiring Norwegian vocabulary.

The finding is also consistent with the full transfer hypothesis (Schwartz & Sprouse, 1996). However, the evidence does not lend full support to the hypothesis of Bardel and Falk (2007), which states that L2 functions as a filter between L1 and L3, completely blocking transfer from L1 to L3. As mentioned, the experiment does not control for the participants' L2. Although all participants spoke English, their proficiency was not asked for or tested in any way. Consequently, it is not possible to predict whether transfer might originate in a common L2, like English, for the German participants.

6.2.3. Word frequency

Balota and Chumbley (1984) found that word frequency is particularly influential on the reaction times of lexical decision tasks. The reaction times to six of the German target words were outliers.

However, the baseline was designed to even out the lexical effects (see section 4.5.2).

6.3. Vocabulary test results

There was no correlation between the vocabulary test results and the priming effects. The findings indicate that the participants' explicit and declarative knowledge of Norwegian is no indication of implicit and procedural knowledge. Ellis (1993), categorized four types of language knowledge based the interaction between implicit/explicit knowledge and declarative/procedural knowledge, explicit and declarative (Type A), implicit and declarative (Type B), explicit and procedural (Type C) and implicit and procedural (Type D).

The Spanish L1 speakers and the German L1 speakers performed equally well on the vocabulary test, suggesting that the Spanish L1 speakers had reached the same level of type A lexical knowledge of Norwegian as the German participants. According to Ellis (1993), type A knowledge is what is taught in classes, while Type D knowledge is native-like fluency in a language. Type D knowledge was investigated through the priming effects in the RT experiment. The German L1 speakers experienced priming effects from meaning related prime words, indicating that they had reached type D knowledge of Norwegian lexicon. The Spanish L1 speakers, however had still not reached this type of knowledge, although they were at the same level of type A knowledge.

Sonbul and Schmitt (2013) have investigated the relationship between implicit and explicit lexical knowledge. Three different tests were used to measure explicit and implicit knowledge. Form recall and form recognition tests were used to measure explicit knowledge, and a priming test to measure implicit knowledge. They found that there is a "clear disassociation between implicit and explicit lexical knowledge, and [that] one does not imply the other" (Sonbul & Schmitt, 2013, p. 151).

The two lowest scoring participants were among those who had not received any instruction in the Norwegian language. This suggests that explicit knowledge, which was tested in the vocabulary test, is something that can be taught and learnt in a class.

As mentioned above, there was no correlation between priming effects and vocabulary test scores. However, a significant positive correlation was detected between the reaction times and the vocabulary test scores of the German participants. The high scorers of the vocabulary test responded faster throughout the RT experiment, indicating that a higher explicit and declarative knowledge of Norwegian facilitates reaction to Norwegian-German cross-linguistic stimuli.

It is possible that the German high scorers' decrease in RTs makes it more difficult to detect priming effects for those participants. As reaction times are already low, a possible further decrease could be smaller and difficult to detect.

6.3.1. Relevance to teaching

The results indicate that German L1 speakers learning Norwegian might benefit from the language similarities between German and Norwegian. However, the advantage is only detectable in online testing. In the offline vocabulary test, the two groups perform equally well.

Although the results did not indicate any advantage for the German L1 speakers over the Spanish L1 speakers in explicit knowledge, the results suggest that second language learners could benefit from being taught alongside learners with the same native language. If implicit knowledge increases faster for the German L1 speakers than for the Spanish L1 speakers, they might progress faster in explicit knowledge as well if the classes were given separately. Further research is necessary to investigate these possibilities.

If structures in L1 can predict the ease and effectiveness with which an L2 is learned – whether transfer is constricted by the processability hierarchy (Pienemann & Keßler, 2011) or not – language learners with different L1 might progress at different speeds. In classes teaching second language, a division of the students based on L1 could be advantageous for learning.

6.4. Further research

For further research, investigation of possible explanations for the different results of Stremme et al. (2014) and the present study would be interesting. One way of investigating these differences could be to still control for native language, but additionally control for L2 and use L2 as the target language. This would yield insight into whether Bardel and Falk's (2007) assumption that L3 is more closely connected to L2 than to L1 might be the cause of the smaller effects of the present study.

Controlling for L2 and possible L3, L4 etc. could also provide further insight into the structure of the multilingual lexicon. This could be done in combination with either a L2 target language experiment or a L1 target language experiment.

It would be interesting to perform a masked priming replication experiment to investigate whether and how this would change the priming effects for meaning and form related word pairs.

In any further study, it would be advantageous to control for background noise. As the present work could not determine the exact effects of the condition of the ventilator, it would have been easier to examine the results if all of the participants had performed the experiment under the same conditions.

As Norwegian, German and Spanish are all Indo-European languages, they do share many linguistic traits. Performing this experiment with languages that are not as closely related could possibly yield more insight into the different hypotheses on transfer.

Another issue to further investigate is the differences between explicit and implicit lexical knowledge. As no correlation was found between the vocabulary test scores and the priming effects, although the same stimuli was used in both tests, it seems that there are different underlying processes for performing the two different tasks.

It could also be interesting to design studies that investigate constructions rather than lexemes, to further examine the hypotheses of transfer. For this purpose, a self-paced reading test could be used. This is a method used by fellow student Eli Rugaard (2016) in her master thesis.

Chapter 7

Conclusion

The results from the RT experiment show that German L1 speakers experience a significant priming effect (-19 ms) from meaning related Norwegian prime words. The detected effects are small compared to effects reported in the pilot study (Stremme et al., 2014). For the Spanish L1 speakers, there is no such priming effect. This finding suggests that a language learner's L1 might affect the acquisition of a given L3. The influence of L1 on L3 lexical acquisition could be due to positive transfer.

The difference in effect size between the pilot study and the present study could be a result of the difference in experiment design. If priming from L3 to L2 produces a greater priming effect than priming from L3 to L1, it supports the claim that L2 and L3 is more strongly connected than L1 and L3, but not that L2 works as a filter between L1 and L3, completely blocking the L3 from L1 influence (Bardel & Falk, 2007).

The vocabulary test scores were almost equally distributed in the two language groups. The two lowest scoring participants were among those who had not received any instruction in the Norwegian language. This suggests that explicit knowledge, which was tested in the vocabulary test, is something that can be taught and learnt.

The vocabulary test scores did not correlate with the priming effects from RT experiment. However, for the German participants, there was a correlation between higher vocabulary test scores and lower reaction times in general. This suggests that declarative knowledge of a language could decrease the reaction time to stimuli in that language. The decrease general in reaction times could have made priming effects more difficult to detect.

The fact that there was no correlation between priming effects and vocabulary test scores points to a division between implicit and explicit knowledge. While the German L1 speakers seem to have reached type D knowledge (Ellis, 1993), the Spanish L1 speakers have not.

The results of the study could have implications for second language teaching. Language background has been shown to affect the speed of progress for fluency in a language, and this supports a possibility to adapt language instruction to the language background of the learners. Exactly how this is optimally done is a subject worthy of more research.

Appendix A. Spanish word list

Cognates

Part 1

Wnr	*Norwegian*	*Spanish*
sw1	sjåfør	CHÓFER
sw2	baby	BEBÉ
sw3	turist	TURISTA
sw4	doktor	DOCTOR
sw5	rotte	RATA
sw6	mamma	MAMÁ
sw7	tiger	TIGRE
sw8	liste	LISTA
sw9	tomat	TOMATE
sw10	dusj	DUCHA
sw11	frukt	FRUTA
sw12	pære	PERA

Part 2

Wnr	*Norwegian*	*Spanish*
sw9	tomat	TOMATE
sw10	dusj	DUCHA
sw11	frukt	FRUTA
sw12	pære	PERA
sw13	rose	ROSA
sw14	kyst	COSTA
sw15	vin	VINO
sw16	tobakk	TABACO
sw17	papir	PAPEL
sw18	lys	LUZ
sw19	verb	VERBO
sw20	idé	IDEA

Translations

Part 1

Wnr	*Norwegian*	*Spanish*
sw21	okse	TORO
sw22	kvinne	MUJER
sw23	bjørn	OSO
sw24	bil	COCHE
sw25	koffert	MALETA
sw26	ost	QUESO
sw27	pose	BOLSA
sw28	seng	CAMA
sw29	jordbær	FRESA
sw30	dråpe	GOTA
sw31	finger	DEDO
sw32	honning	MIEL

Part 2

Wnr	*Norwegian*	*Spanish*
sw29	jordbær	FRESA
sw30	dråpe	GOTA
sw31	finger	DEDO
sw32	honning	MIEL
sw33	klokke	RELOJ
sw34	innsjø	LAGO
sw35	negl	UÑA
sw36	stol	SILLA
sw37	snø	NIEVE
sw38	fot	PIE
sw39	varme	CALOR
sw40	reise	VIAJE

Unrelated

Part 1

Wnr	*Norwegian*	*Spanish*
sw41	konge	PERRO
sw42	lærer	ARAÑA
sw43	fugl	TÍO
sw44	venn	CERDO
sw45	geit	NOVIO
sw46	skap	FLOR
sw47	hode	CASA
sw48	skjerf	MONTAÑA
sw49	elv	MUNDO
sw50	briller	ÁRBOL
sw51	saks	FALDA
sw52	ski	LLAVE

Part 2

Wnr	*Norwegian*	*Spanish*
sw49	elv	MUNDO
sw50	briller	ÁRBOL
sw51	saks	FALDA
sw52	ski	LLAVE
sw53	hage	LIBRO
sw54	kopp	DINERO
sw55	melding	AIRE
sw56	vindu	POSTRE
sw57	sko	OREJA
sw58	høyde	VIDA
sw59	sinne	PALABRA
sw60	fart	AÑO

False cognates

Part 1

Wnr	*Norwegian*	*Spanish*
sw61	panne	PAN
sw62	åre	ORO
sw63	plate	PLATA
sw64	horn	HORNO
sw65	sene	CENA
sw66	kart	CARTA
sw67	buffet	BUFETE
sw68	nummer	NOMBRE
sw69	suksess	SUCESO
sw70	trening	TREN
sw71	notis	NOTICIA
sw72	rop	ROPA

Part 2

Wnr	*Norwegian*	*Spanish*
sw69	suksess	SUCESO
sw70	trening	TREN
sw71	notis	NOTICIA
sw72	rop	ROPA
sw73	tramp	TRAMPA
sw74	mann	MANO
sw75	bot	BOTA
sw76	bur	BURRO
sw77	bod	BODA
sw78	kode	CODO
sw79	gåte	GATO
sw80	måne	MONO

Appendix B. German word list

Cognates

Part 1			Part 2		
Wnr	***Norwegian***	***German***	***Wnr***	***Norwegian***	***German***
gw1	katt	KATZE	gw9	hus	HAUS
gw2	lærer	LEHRER	gw10	flaske	FLASCHE
gw3	bror	BRUDER	gw11	fabrikk	FABRIK
gw4	atlet	ATHLET	gw12	nese	NASE
gw5	fetter	VETTER	gw13	brød	BROT
gw6	dato	DATUM	gw14	sag	SÄGE
gw7	bok	BUCH	gw15	kne	KNIE
gw8	resept	REZEPT	gw16	glass	GLAS
gw9	hus	HAUS	gw17	arbeid	ARBEIT
gw10	flaske	FLASCHE	gw18	økonomi	ÖKONOMIE
gw11	fabrikk	FABRIK	gw19	pris	PREIS
gw12	nese	NASE	gw20	råd	RAT

Translations

Part 1			Part 2		
Wnr	***Norwegian***	***German***	***Wnr***	***Norwegain***	***German***
gw21	mark	WURM	gw29	seng	BETT
gw22	hest	PFERD	gw30	hevn	RACHE
gw23	bonde	BAUER	gw31	fløte	SAHNE
gw24	kone	EHEFRAU	gw32	saks	SCHERE
gw25	bil	AUTO	gw33	vaksine	IMPFSTOFF
gw26	mat	ESSEN	gw34	nøkkel	SCHLÜSSEL
gw27	ost	KÄSE	gw35	tommel	DAUMEN
gw28	verden	WELT	gw36	blyant	BLEISTIFT
gw29	seng	BETT	gw37	hjørne	ECKE
gw30	hevn	RACHE	gw38	bukse	HOSE
gw31	fløte	SAHNE	gw39	skje	LÖFFEL
gw32	saks	SCHERE	gw40	hav	MEER

Unrelated

Part 1			**Part 2**		
Wnr	***Norwegian***	***German***	***Wnr***	***Norwegian***	***German***
gw41	rotte	JAHR	gw49	okse	GARTEN
gw42	rev	SCHNEIDER	gw50	tid	ZORN
gw43	skap	BAUM	gw51	sokk	UHR
gw44	øre	SESSEL	gw52	spill	ERBSE
gw45	fjell	KISSEN	gw53	bord	GABEL
gw46	snø	HANDY	gw54	avis	SCHAL
gw47	finger	PULLI	gw55	sofa	FENSTER
gw48	løk	BAHNHOF	gw56	trikk	FLUSS
gw49	okse	GARTEN	gw57	ord	HÖHE
gw50	tid	ZORN	gw58	knapp	BILD
gw51	sokk	UHR	gw59	varme	RECHNER
gw52	spill	ERBSE	gw60	sko	FAHRRAD

False cognates

Part 1			**Part 2**		
Wnr	***Norwegian***	***German***	***Wnr***	***Norwegian***	***German***
gw61	vaffel	WAFFE	gw69	kinn	KIND
gw62	enke	ENKEL	gw70	mat	MADE
gw63	vann	WAND	gw71	valg	WALD
gw64	kopp	KOPF	gw72	trakt	TRACHT
gw65	krutt	KRAUT	gw73	glede	KLEID
gw66	urt	ORT	gw74	hjem	HEMD
gw67	vorte	WURST	gw75	eim	EIMER
gw68	bær	BÄR	gw76	innside	INSEL
gw69	kinn	KIND	gw77	løv	LÖWE
gw70	mat	MADE	gw78	bod	BODEN
gw71	valg	WALD	gw79	hassel	HASE
gw72	trakt	TRACHT	gw80	møll	MÜLL

Appendix C. Questionnaire

PARTICIPANT NUMBER:

Native language:	☐ German	☐ Spanish

Age:	☐ under 18	☐ 18-25	☐ 25-35	☐ 35-45	☐ 45-55	☐ over 55
Gender:	☐ Male	☐ Female				

Which other languages do you speak – ranked by proficiency from best to worst?

Your level of Norwegian:

Writing:	1	2	3	4	5	6
Reading:	1	2	3	4	5	6
Speaking:	1	2	3	4	5	6

Your level of your native language:

Writing:	1	2	3	4	5	6
Reading:	1	2	3	4	5	6
Speaking:	1	2	3	4	5	6

Appendices

Bibliography

Albert, M. L., & Obler, L. K. (1978). *The bilingual brain : neuropsychological and neurolinguistic aspects of bilingualism*. New York: Academic Press.

Ard, J., & Homburg, T. (1992). Verification of Language Transfer. In S. Gass & L. Selinker (Eds.), *Language Transfer in Language Learning*: John Benjamins Publishing company.

Baayen, R. H. (2014). Experimental and psycholinguistic approaches to studying derivation. doi:10.15496/publikation-4877

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-Effects Modeling with Crossed Random Effects for Subjects and Items. *Journal of Memory and Language, 59*(4), 390-412. doi:10.1016/j.jml.2007.12.005

Baayen, R. H., & Milin, P. (2010). Analyzing Reaction Times. *International Journal of Psychological Research, 3*(2), 12-28.

Balota, D. A., & Chumbley, J. I. (1984). Are lexical decisions a good measure of lexical access? The role of word frequency in the neglected decision stage. *Journal of experimental psychology. Human perception and performance, 10*(3), 340.

Bardel, C., & Falk, Y. (2007). The Role of the Second Language in Third Language Acquisition: The Case of Germanic Syntax. *Second Language Research, 23*(4), 459-484. doi:10.1177/0267658307080557

Bargh, J. (2014). The historical origins of priming as the preparation of behavioral responses: Unconscious carryover and contextual influences of real-word importance. *Soc. Cogn., 32*, 209-224.

Bauvillain, C., & Grainger, J. (1987). Accessing Interlexical Homographs: Some Limitations of a Language-Selective Access. *Journal of Memory and Language*(26), 658-672.

Birdsong, D. (2006). Age and Second Language Acquisition and Processing: A Selective Overview. *Language Learning, 56*(s1), 9.

Braun, M., Jacobs, A. M., Hahne, A., Ricker, B., Hofmann, M., & Hutzler, F. (2006). Model-generated lexical activity predicts graded ERP amplitudes in lexical decision. *Brain Research, 1073*, 431-439. doi:10.1016/j.brainres.2005.12.078

Brysbaert, M., & Duyck, W. (2010). Is it time to leave behind the Revised Hierarchical Model of bilingual language processing after fifteen years of service? *BILINGUALISM-LANGUAGE AND COGNITION*. doi:10.1017/S1366728909990344

Buchstaller, I., & Khattab, G. (2013). Population Samples. In R. J. Podesva & D. Sharma (Eds.), *Research methods in linguistics* (pp. 74-95). Cambridge: Cambridge University Press.

Carroll, S. E. (1992). On Cognates. *Second Language Research, 8*(2), 93-119.

Cedrus. (2015). Response Box RB-540. Retrieved from http://cedrus.com/rb_series/

Cenoz, J. (2001). The Effect of Linguistic Distance, L2 status and Age on Cross-Linguistic Influence in Third Language Acquisition. In J. Cenoz, B. Hufeisen, & U. Jessner (Eds.), *Cross-Linguistic Influence in Third Language Acquisition: Psycholinguistic Perspectives* (pp. 8-20). Clevedon: Multilingual Matters.

Cenoz, J., Hufeisen, B., & Jessner, U. (2001). Introduction. In J. Cenoz, B. Hufeisen, & U. Jessner (Eds.), *Cross-Linguistic Influence in Third Language Aquisition : Psycholinguistic Perspectives* (pp. 1-6). Clevedon: Multilingual Matters.

Collins English Dictionary. (2012). Lexical decision task *Collins English Dictionary - Complete & Unabridged Digital Edition*: HarperCollins Publishers.

Croft, W. (2003). *Typology and universals* (2nd ed. ed.). Cambridge: Cambridge University Press.

de Groot, A. M. B. (2013). Reading. In F. Grosjean & P. Li (Eds.), *The Psycholinguistics of Bilingualism*. Oxford, UK: Wiley-Blackwell.

DeKeyser, R. M. (2012). Interactions between Individual Differences, Treatments, and Structures in SLA. *Language Learning: A Journal of Research in Language Studies, 62*(2 [Supplement]), 189-200. doi:10.1111/j.1467-9922.2012.00712.x

DeKeyser, R. M. (2013). Age Effects in Second Language Learning: Stepping Stones toward Better Understanding. *Language Learning, 63*(1), 52-67. doi:10.1111/j.1467-9922.2012.00737.x

Dijkstra, T., Timmermans, M., & Schriefers, H. (2000). On Being Blinded by Your Other Language: Effects of Task Demands on Interlingual Homograph Recognition. *Journal of Memory and Language, 42*(4), 445-464. doi:10.1006/jmla.1999.2697

Dryer, M. S., & Haspelmath, M. (Eds.). (2013). *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology.

Ellis, R. (1993). The Structural Syllabus and Second Language Acquisition. *TESOL Quarterly, 27*(1), 91-113. doi:10.2307/3586953

Encyclopædia Britannica. (2015). "Romance languages".

Falk, Y., & Bardel, C. (2011). Object Pronouns in German L3 Syntax: Evidence for the L2 Status Factor. *Second Language Research, 27*(1), 59-82. doi:10.1177/0267658310386647

Field, F. W. (2011). *Key concepts in bilingualism*. Basingstoke: Palgrave Macmillan.

Forster, K. (2015). Masked Priming. Retrieved from http://www.u.arizona.edu/~kforster/priming/

Gabryś-Barker, D. (2005). *Aspects of Multilingual Storage, Processing and Retrieval*. Katowice: Wydawnictwo Univerytetu Śląskiego.

Grosjean, F. (2013). Bilingualism: A Short Introduction. In F. Grosjean & P. Li (Eds.), *The Psycholinguistics of Bilingualism*. Oxford, UK: Wiley-Blackwell.

Gullberg, M., Roberts, L., Dimroth, C., Veroude, K., & Indefrey, P. (2010). Adult Language Learning After Minimal Exposure to an Unknown Natural Language. *Lang. Learn., 60*, 5-24.

Haberzettl, S. (2005). *Der Erwerb der Verbstellungsregeln in der Zweitsprache Deutsch durch Kinder mit russischer und türkischer Muttersprache* Linguistische Arbeiten (Max Niemeyer Verlag), Vol. 495.

Hawkins, J. A. (1983). *Word Order Universals*. San Diego: Academic Press.

Hell, J. G. v., & Dijkstra, T. (2002). Foreign Language Knowledge Can Influence Native Language Performance in Exclusively Native Contexts. *Psychonomic Bulletin & Review: A Journal of the Psychonomic Society, Inc., 9*(4), 780-789. doi:10.3758/BF03196335

Hickey, R. (2012). *The Handbook of Language Contact*. Hoboken: Wiley.

Hyltenstam, K., & Abrahamsson, N. (2003). Maturational constraints in secondlanguage acquisition. In C. J. Doughty & M. H. Long (Eds.), *Handbook of Second Language Acquisition* (pp. 539-588). Oxford, UK: Blackwell.

Håkansson, G., Pienemann, M., & Sayehli, S. (2002). Transfer and typological proximity in the context of L2 processing. *Second Language Research, 18*(3), 250-273.

Indefrey, P., & Gullberg, M. (2010). The Earliest Stages of Language Learning: Introduction. *Lang. Learn., 60*, 1-4.

Johnson, J. S., & Newport, E. L. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology, 21*(1), 60-99. doi:10.1016/0010-0285(89)90003-0

Kaiser, E. (2013). Experimental Paradigms in Psycholinguistics. In R. J. Podesva & D. Sharma (Eds.), *Research Methods in Linguistics* (pp. 135-168). New York: Cambridge University Press.

Kim, H.-S. (1987). Typological approach to language universals and language transfer in second language acquisition: Language typology of word order: ProQuest Dissertations Publishing.

Kjellberg, A., & Landström, U. (1994). Noise in the office: Part II — The scientific basis (knowledge base) for the guide. *International Journal of Industrial Ergonomics, 14*(1), 93-118. doi:10.1016/0169-8141(94)90008-6

Kohfeld, D. L., & Goedecke, D. W. (2013). Intensity and predictability of background noise as determinants of simple reaction time. *Bulletin of the Psychonomic Society, 12*(2), 129-132. doi:10.3758/bf03329648

Kolers, P. A. (1963). Interlingual word associations. *Journal of Verbal Learning and Verbal Behavior, 2*(4), 291-300. doi:http://dx.doi.org/10.1016/S0022-5371(63)80097-3

Kroll, J. F., Gerfen, C., & Dussias, P. E. (2010). Laboratory Designs and Paradigms. In L. Wei & M. G. Moyer (Eds.), *Research Methods in Bilingualism and Multilingualism* (3 ed.). Oxford, UK: Blackwell.

Kroll, J. F., & Stewart, E. (1994). Category Interference in Translation and Picture Naming: Evidence for Asymmetric Connections Between Bilingual Memory Representations. *Journal of Memory and Language, 33*(2), 149-174. doi:10.1006/jmla.1994.1008

Kroll, J. F., van Hell, J. G., Tokowicz, N., & Green, D. W. (2010). The Revised Hierarchical Model: A critical review and assessment. *Bilingualism (Cambridge, England), 13*(3), 373-381. doi:10.1017/S136672891000009X

Lehrer, J. (2010). The Truth Wears Off. *The New Yorker, 86*(40), 52.

Lemhöfer, K., & Dijkstra, T. (2004). Recognizing Cognates and Interlingual Homographs: Effects of Code Similarity in Language-Specific and Generalized Lexical Decision. *Memory & Cognition, 32*(4), 533-550. doi:10.3758/BF03195845

Lenneberg, E. H. (1967). *Biological foundations of language*. New York: Wiley.

Levelt, W. J. M. (1989). *Speaking : from intention to articulation*. Cambridge, Mass: MIT Press.

Levelt, W. J. M. (2001). Spoken Word Production: A Theory of Lexical Access. *Proceedings of the National Academy of Sciences of the United States of America, 98*(23), 13464-13471.

Li, P. (2013). Successive Language Acquisition. In F. Grosjean & P. Li (Eds.), *The Psycholinguistics of Bilingualism*. Oxford, UK: Wiley-Blackwell.

Luk, G., & Bialystok, E. (2013). Bilingualism is not a categorical variable: Interaction between language proficiency and usage. doi:10.1080/20445911.2013.795574

Meyer, D. E., Schvaneveldt, R. W., & Grant, D. A. (1971). Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology, 90*(2), 227-234. doi:10.1037/h0031564

Monner, D., Vatz, K., Morini, G., Hwang, S.-O., & DeKeyser, R. (2013). A Neural Network Model of the Effects of Entrenchment and Memory Development on Grammatical Gender

Learning. *Bilingualism: Language and Cognition, 16*(2), 246-265. doi:10.1017/S1366728912000454

Perani, D., Paulesu, E., Galles, N., Dupoux, E., Dehaene, S., Bettinardi, V., . . . Mehler, J. (1998). The bilingual brain - Proficiency and age of acquisition of the second language. *Brain, 121*, 1841-1852.

Pienemann, M. (1998). *Language Processing and Second Language Development: Processability theory*. Amsterdam: Amsterdam : John Benjamins Publishing Company.

Pienemann, M., & Keßler, J.-U. (2011). *Studying Processability Theory: An Introductory Textbook*. Amsterda: Amsterda : John Benjamins Publishing Company.

Pienemann, M., Keßler, J., U., & Itani-Adams, Y. (2011). Comparing levels of processability across languages (Vol. 15, pp. 128-146).

Potter, M. C., So, K.-F., Eckardt, B. V., & Feldman, L. B. (1984). Lexical and conceptual representation in beginning and proficient bilinguals. *Journal of Verbal Learning and Verbal Behavior, 23*(1), 23-38. doi:10.1016/S0022-5371(84)90489-4

R Development Core Team. (2008). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from http://www.R-project.org

Rugaard, E. (2016). *Balancing Bilingualism down the Garden Path. (Unpublished master thesis)*. Universitetet i Bergen. Bergen.

Sandøy, H. (2000). *Lånte fjører eller bunad? : om importord i norsk* (Vol. nr 136). Oslo: Landslaget for norskundervisning Cappelen akademisk forl.

Schwartz, B. D., & Sprouse, R. A. (1996). L2 Cognitive States and the Full Transfer/Full Access Model. *Second Language Research, 12*(1), 40-72. doi:10.1177/026765839601200103

Singleton, D. (1999). *Exploring the second language mental lexicon*. Cambridge: Cambridge University Press.

Singleton, D., & Ryan, L. (2004). *Language Acquisition : The Age Factor* (2nd ed. ed.). Clevedon: Multilingual Matters.

Sonbul, S., & Schmitt, N. (2013). Explicit and Implicit Lexical Knowledge: Acquisition of Collocations under Different Input Conditions. *Language Learning, 63*(1), 121-159. doi:10.1111/j.1467-9922.2012.00730.x

Store norske leksikon. (2009). *Arveord*.

Stremme, C. L. (2015). *A study of cross-linguistic form and meaning priming through reaction times in lexical decision tasks.* (Master thesis), Universitetet i Bergen, Bergen. Retrieved from http://bora.uib.no/bitstream/handle/1956/10874/141131953.pdf?sequence=1

Stremme, C. L., Troland, V., & Johansson, C. (2014). *Meaning advantage in cross-linguistic priming.* Paper presented at the Psycholinguistics in Flanders, Oostende.

Szubko-Sitarek, W. (2014). *Multilingual Lexical Recognition in the Mental Lexicon of Third Language Users*. Berlin: Springer Berlin Heidelberg.

Söderlund, G., Sikström, S., Loftesnes, J., & Sonuga-Barke, E. (2010). The effects of background white noise on memory performance in inattentive school children. *Behavioral and Brain Functions, 6*, 55. doi:10.1186/1744-9081-6-55

Trimmel, M., & Poelzl, G. (2006). Impact of background noise on reaction time and brain DC potential changes of VDT-based spatial attention. *Ergonomics, 49*(2), 202-208. doi:10.1080/00140130500434986

Tymczyńska, M. (2012). Cross-linguistic Influences in Multilingual Language Acquisition. In D. Gabrys-Barker (Ed.), *Second Language Learning and Teaching*. Berlin, Heidelberg: Springer Berlin Heidelberg, Berlin, Heidelberg.

Van Kesteren, R., Dijkstra, T., & De Smedt, K. (2012). Markedness effects in Norwegian–English bilinguals: Task-dependent use of language-specific letters and bigrams. *The Quarterly Journal of Experimental Psychology, 65*(11), 2129-2154. doi:10.1080/17470218.2012.679946

Weinreich, U. (1953). *Languages in contact : findings and problems* (Vol. no. 1). New York: Linguistic Circle of New York.

Wilkinson, D., Nicholls, S., Pattenden, C., Kilduff, P., & Milberg, W. (2008). Galvanic vestibular stimulation speeds visual memory recall. *Experimental Brain Research, 189*(2), 243-248. doi:10.1007/s00221-008-1463-0

Winter, B. (2013). Linear Models and linear mixed effects models in R with linguistic applications. arXiv:1308.5499.